Evolutionary Algorithms and Frequent Itemset Mining for Analyzing Epileptic Oscillations

A Dissertation
Presented to
The Academic Faculty
By

Otis Lkuwamy Smart

In Partial Fulfillment
Of the Requirements for the Degree:
Doctor of Philosophy in Electrical Engineering

Georgia Institute of Technology

May 2007

Evolutionary Algorithms and Frequent Itemset Mining for Analyzing Epileptic Oscillations

Approved by:

Dr. George J. Vachtsevanos, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Brian Litt, M.D., Co-Advisor
Department of Neurology and Bioengineering
Hospital of the University of Pennsylvania
University of Pennsylvania

Dr. Douglas B. Williams
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Ayanna M. Howard
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Robert J. Butera
School of Electrical and Computer Engineering
Interdisciplinary Bioengineering Graduate Program
Georgia Institute of Technology

Dr. Javier Echauz
BioQuantix Corporation

Date Approved: 8 March 2007

Dedication

This milestone is dedicated to current and future African American Ph.D. students whom I encourage to know the rules of the game and harness their God-given brain while persevering through the pain to accomplish the same.

Acknowledgements

I give utmost thanks to God for always providing the talent, experience, diligence, perseverance, resilience, success, failure, support, and love to achieve this goal. Many different people and experiences have influenced the completion of this work—with You being the constant. I thank You for Your blessings, including the following supporters.

I am appreciative of my co-advisors Dr. George Vachtsevanos and Dr. Brian Litt for their conviction, supervision, evaluation and point of view throughout this project. I am grateful to the collaborators at the University of Pennsylvania, Children's Hospital of Philadelphia, and the Mayo Clinic in Rochester, MN for their contributions to this research. In particular, Dr. Litt, Dr. Gregory Worrell, Dr. Dennis Dlugos, and Dr. Eric Marsh for their time in examining data, sharing medical expertise and discussing results. I am grateful to Dr. Javier Echauz, Dr. Hiram Firpi, Dr. Andrew Gardner, Steve Cranstoun, and Lauren Burrell for imparting technical expertise in the development of this work. I express gratitude to the members of the committee, Dr. Vachtsevanos, Dr. Litt, Dr. Douglas Williams, Dr. Ayanna Howard, Dr. Robert Butera, and Dr. Echauz, for their time and efforts in reviewing this thesis and convening for the defense.

I value the examples embodied in Dr. Gary May, Dr. Mark Smith, Dr. Keith Oden, Dr. Howard, Dr. Freeman Rufus, Dr. Raheem Beyah, Dr. Terrell Neal, Dr. Amos Johnson, Dr. Robert Bozeman, Dr. Robert Dixon, Dr. Abdelkrim Brania, Dr. Dalinda Brown-Clark, Dr. Melissa Green, Dr. Sonya Greene, Clarence Brown, Dr. Harriet Jenkins, Dr. Norma J. Burgess, and Charlie Burgess; insight instilled by Suzette

Willingham, Jacqueline Trappier, Sharon Lawrence, Jill Auerbach, Dr. David Hertling, Dr. Bonnie Ferri, Marylou Mycko, Jean Pierre Collier, Reginald Rowell, Robert and Stella Wilson, Harry and Mattie Williams, Mr. And Mrs. Burgess, Luverta Smart, Otis Lee Smart Jr., and Viola Manigo-Smart; and faithful friendships with Sherron Burgess, Joseph Walton III, Jim Fraction, Dave Williams, Lauren Burrell, Dr. Firpi, Dr. Neal, Gerald DeJean, Brett Matthews, Chris Green, Sekou Remy, Jacqueline Fairley, Karolyn Babalola, Jamal Wilson, Clarence Wardell III, Russell Marzette, and Rodney Sawyer that cultivated my growth in and completion of the Ph.D. program at Georgia Tech.

I am indebted to the Harriet G. Jenkins Predoctoral Fellowship Program, Southern Regional Educational Board Doctoral Scholars Program, Opportunity Scholars Program, Georgia Tech Facilitating Academic Careers in Engineering & Science, Georgia Engineering Foundation, Georgia Regents, General Electric Faculty for the Future Program, my co-advisors, and Dr. McClellan for funding my education in graduate school as well as Dr. Vachtsevanos, Dr. Litt, Dr. McClellan, Dr. May, Dr. Oden, Jill Auerbach, Dr. Dixon, Dr. Bozeman, Sharon Lawrence and C. Green for their assistance in obtaining the funding.

Finally, I am thankful to my family, especially my brothers Mike and Jason, for their encouragements and understanding despite times that I missed or abbreviated moments with them to finish this dissertation. I am eternally thankful for my brothers and sisters in Christ who have prayed for Your will through me. I am thankful to God for my fiancé, Sherron Lynn Burgess, who truly is my biggest fan: I will always remember your unconditional endurance, sacrifice, love, and support during this painful process and thank God for your special role in my life.

Table of Contents

List of Tables

List of Figures

List of Keywords and Abbreviations

intracranial electroencephalogram (IEEG)

electroencephalogram (EEG)

epilepsy monitoring unit (EMU)

earliest electrographic change (EEC)

unequivocal EEG onset (UEO)

ictal

interictal

preictal

epileptic oscillation

epileptic high frequency oscillation (EHFO)

high frequency epileptiform oscillation (HFEO)

seizure onset zone (SOZ)

ictal onset zone (IOZ)

epileptogenic zone (EZ)

evolutionary algorithm (EA)

genetic programming (GP)

particle swarm optimization (PSO)

genetic programming of classical features (GP-CF)

genetic programming of delayed data (GP-DD)

particle swarm optimization of a neural network (PSO-NN)

artificial features (a-features)

classical features or conventional features (c-features)

feature series

malleable feature

finite impulse response (FIR)

frequent itemset mining (FIM)

analysis of variance (ANOVA)

sensitivity (SENS)

specificity (SPEC)

true positive (TP)

false positive (FP)

true negative (TN)

false negative (FN)

receiver operating characteristic (ROC) curve

area underneath the receiver operating characteristic (AUROC) curve

area underneath the [ROC] curve (AUC)

Fisher's discriminant ratio (FDR)

Probability density (distribution) function (PDF)

Probability mass function (PMF)

Hidden Markov Model (HMM)

Markov Model (MM)

Hitting time

First hitting time

Brownian motion

Brownian particle

Summary

The objective of this research is to develop an automated methodology that maps epileptic networks in patients with epileptic seizures. Epilepsy afflicts over 60 million people worldwide with limited options for treatment, one-third of who have seizures that cannot be controlled by antiepileptic drugs (AED's). For 7% of patients with epilepsy, surgery is an option because their seizures reliably begin in one discrete region. However, effective surgical procedures typically require removing a significant volume of brain tissue due to the lack of a reliable method for pinpointing the exact location of seizure onset and the pathways through which seizures are generated and spread. Consequently, medical or surgical treatment cannot control or remedy epilepsy in over 15 million individuals with the disorder, partly due to poorly localized or multi-focal seizure onsets. Although new technologies based upon pre-emptive electrical stimulation, the infusion of drugs, or focal cooling may be viable alternatives to surgical resection for patients with intractable seizures, these approaches still require a reliable means to map epileptic networks and determine the best sites for therapeutic intervention.

It has been suggested that certain epileptic high frequency oscillations within the intracranial electroencephalogram (IEEG) both between seizures (interictal) and preceding a seizure (pre-ictal) may localize regions important to seizure generation on the IEEG. The work below aims to map epileptogenic tissue by detecting epileptic oscillations in the IEEG and statistically identifying regions of brain in which the oscillations are concentrated. The methodology to accomplish this aim incorporates two primary components: 1) a module that detects the oscillations using an feature via

evolutionary algorithms; and 2) a module that statistically filters the output of the first module using *frequent itemset mining* (FIM), which discovers chronic patterns across electrodes. Together, these components provide decision-support for an expert to map networks in the brain that generate seizures and guide a suitable therapy for patients.

Prior attempts to accurately detect similar electrophysiological events in the IEEG have exhibited marginal performance. This may be because the methods lack robust choices for the selection, extraction, and combination of features that effectively distinguish epileptic events from non-epileptic events. This research successfully implements evolutionary algorithms (i.e., *genetic programming*, *particle swarm optimization*) to detect epileptic oscillations in data with very poor signal-to-noise ratio (SNR)—even after filtering—and a low bandwidth up to 100 Hz.

The presented detector achieves median values near and above 80% sensitivity and 85% specificity over a sample of testing records consisting of 10-16 records with at least 30 epileptic oscillations per record across six subjects. An analysis of variance (ANOVA) reveals that the presented method outperforms (in terms of sensitivity and specificity) a naïve method, which represents a "standard," to detect epileptic oscillations using an energy feature and a threshold. Permutation testing demonstrates that the method incorporating features returned by an evolutionary algorithm (EA) does not perform equal to random chance at P-values near zero. The FIM clusters electrodes with high concentrations of epileptic oscillations, where most couplings overlap with areas of ictal (seizure) onset and epileptogenic zones defined by clinically by expert physicians.

These findings and results are significant, considering that this particular field of study in epilepsy is not established, meaning that this research represents some of the first

efforts to investigate that epileptic high frequency oscillations and that no definitive benchmarks have been established. Optimistically, this work provides a benchmark for detecting and analyzing epileptic oscillations. This research contributes an efficient means to create and fuse quality features, techniques to evaluate the quality of a feature, and an improved method to detect abnormal physiological events key to reliably mapping regions of dysfunctional brain that constitute epileptic networks. The ultimate goal is to use this new knowledge to understand the generation of seizures in these individuals and disrupt the mechanism. Hopefully, this research will lead to a better control of seizures and an improved quality of life for the millions of persons affected by epilepsy.

Chapter 1: Introduction

1.1 Motivation

Over 60 million people have epilepsy, a neurological disorder characterized by recurrent seizures that occur without a clear warning [2]. For approximately 66% of the persons affected by this disorder, anti-epileptic drugs (AED's) can control seizures, but side effects are common [2]. Some people with partial (focal) onset seizures, which originate in an unambiguous portion of the brain that is often coincident with lesions, may be candidates for surgery to remove the dysfunctional tissue. More commonly, however, persons have seizures that do not arise from visible lesions, erupt throughout the entire brain at once, and may be poorly localized. For these people with medically refractory seizures, surgery may not be feasible because the disorder manifests in multiple locations or is too complex to be localized.

New therapies, such as implantable medical devices based upon focal cooling, drug-infusion or electrical stimulation, possess strong potential for treating patients with medically refractory seizures and may become a viable option in cases where surgery or other conventional therapies are not apt. However, the efficacy of these and currently available treatments depends upon a reliable means to localize key region(s) of brain that are involved in ictogenesis, or the generation of seizures. That is, key regions must be localized to exactly determine where to treat damaged tissue. Though a very daunting dilemma that challenges current technological limits and existing knowledge of neuroscience, the ability to define and characterize epileptic networks that are involved in

ictogenesis is of increasing interest to researchers, clinicians, and the medical device industry. Study of the electroencephalogram (EEG) is one means to achieve this goal.

In analyzing brain activity from intracranial electrodes, investigators have defined normal and abnormal oscillations. The epilepsy literature focuses on abnormal oscillations that are separated by their dominant frequencies and reportedly generated by epileptic networks. For instance, Bragin and Engle [105-109] refer to ripples (200 Hz) and fast ripples (500 Hz) from microelectrode recordings; Worrell et al. [53, 127] identify lower frequency pathological oscillations (60-100 Hz), referred to as high frequency epileptiform oscillations (HFEO's), in epileptic patients; and Jirsch et al. [163] discover distinct epileptic oscillations localized to areas of seizure onset and spread in depth macro-electrode recordings. These investigations represent converging evidence [53, 105-109, 127, 136, 158, 163] that links a particular pathological oscillation, herein called an epileptic high frequency oscillation (EHFO), within the intracranial electroencephalogram (IEEG) to the localization, generation, and propagation of epileptic seizures.

Consequently, a growing hypothesis in the literature on epilepsy suggests that further investigation of EHFO's can improve our understanding of the dynamics of epilepsy. This improvement in knowledge includes understanding the formation of epileptic networks, which can facilitate better approaches to treat seizures. Developing automated algorithms for mapping these pathological events would contribute significantly to research in epilepsy, given the scientific interest in and potential clinical utility of modulating activity in regions of brain that the algorithms pinpoint. For that purpose, this work focuses on providing a sound, general framework for automatically

mapping epileptic networks according to the occurrence of EHFO's within multiple IEEG recordings.

1.2 Problem Definition

The objective of this research is to develop a comprehensive framework to reliably recognize recurrent locations that generate interictal or preictal epileptic high frequency oscillations in multi-channel recordings of IEEG from patients with epileptic seizures, who are implanted with IEEG electrodes during their evaluation for surgery. Literature strongly suggests that EHFO's may mark dysfunctional areas of brain that constitute epileptic networks, also known as the epileptogenic zone (EZ), or the area of the cortex that is indispensable for the generation of epileptic seizures. This work focuses on the development of a statistically reliable tool for identifying dysfunctional regions and serving as decision-support for a clinician in discerning the epileptogenic zone (EZ).

This problem is approached in two major phases: 1) design a detector of epileptic high frequency oscillations; and 2) implement a subsequent statistical analysis to cluster concentrations of detected epileptic oscillations. Current approaches to detect epileptic oscillations fail to provide a thorough approach for selecting (and combining) features that distinguish the abnormal oscillations from normal oscillations and classifying the feature(s). That is, existing methods typically select a single physics-driven feature—after visual review or mere assumption—that correlates to the energy or amplitude of the event to be detected for feature-extraction, and use a threshold but with an arbitrary setting (e.g. three standard deviations above the mean signal amplitude) for classification. The apparent assumption that "energy features" and a threshold can achieve the best

design for detection holds obvious intuitive and computationally inexpensive appeal. However, to provide a doctor with a reliable tool to help in understanding ictogenesis and treating patients, a more rigorous approach to choosing a feature to extract and the decision in classification is required. Meanwhile, research on the second stage of signal processing is in early development [158], where the below work represents the first attempts to cluster detectable EHFO's for possibly identifying epileptic networks.

Regarding the detector, this work chooses an architecture that fits the traditional modules of pattern classification: signal enhancement, feature extraction, and classification [166]. Moreover, feature extraction is emphasized as the most critical stage in designing a detector [166, 183] for the following reasons:

- If an extracted feature, which serves as input to the classifier, possesses low performance, or low separation between an event of interest and complementary events, then the classifier may poorly discriminate the events.
- Since the computational burden of the classifier depends upon the size and performance of a feature that it processes, it makes more sense to beforehand choose a quality feature up front rather than pass avoidable, complex processing to the classifier.

In particular, *genetic programming* (GP) and *particle swarm optimization* (PSO) are useful techniques for feature selection (and combination) in this category of tasks. The GP and the PSO are automated routines in the family of *evolutionary algorithms* (EA) that can be used to compute an optimal, *artificial feature* for extraction. The feature is optimal in the sense that the heuristics of the EA maximizes an objective function,

which measures the performance of a feature in distinguishing epileptic oscillations from non-epileptic (baseline) activity. The feature is termed *artificial* because the EA returns a computer-crafted result that might not have a physical meaning, as perhaps expected (or desired) by the user of the algorithm.

For instance, the work below elects the application of a GP that permits the computation of a single feature in two possible manners: 1) choosing and combining *physics-based features*, also referred to as *classical features*, or *conventional features*; and 2) creating a *data-driven feature*, or a feature created directly from delays of the data after enhancement of the signals. Additionally, the algorithm for PSO (when coupled with a structure such as a neural network) is used to produce a single data-driven feature. Ultimately, each application of an evolutionary algorithm built subject-specific, artificial features to circumvent the arduous but commonly undertaken tasks of visual review, intuitive derivation, or manual combination of quantities based on understanding the physics of the problem to determine the best physics-based feature(s).

Addressing the second phase of this problem, *frequent itemset mining* (FIM), which is often referred to as *data mining*, is used to complement the detector. To objectively discover a pattern (if any) involving the occurrence of a particular paroxysmal waveform (e.g., HFEO's, fast ripples, sharp waves) within the recordings of long-duration (tens of minutes to hours) multi-channel (nearly 30 to 130 electrodes) IEEG is an impractical task for humans. In fact, this task is even more complex since the waveforms are very brief (milliseconds to a few seconds) in duration and statistical rigor and reliability is required. Frequent itemset mining provides a computationally efficient solution for such a difficult problem. Figure 1.2-1 illustrates the powerful utility of FIM

in clustering electrodes with frequent concentrations of epileptic waveforms. The figure reflects the burden of finding the most informative grouping of channels from many possible combinations—a total of $\sum_{n=1}^{n=32}\binom{32}{n} = 4{,}294{,}967{,}295$ in the case of thirty-two electrodes—of putatively associated channels.

The presented methodology decisively combines techniques in pattern detection and statistical clustering to form an overall pattern recognition system that searches for reproducible patterns among detected occurrences of epileptic waveforms. It is important to note that this methodology does not eliminate human participation. Instead, building a detector using PSO or GP and mining the detections with FIM greatly simplify the problem of studying large volumes of data for possible patterns by a human reviewer. Thus, this project is intended to serve as an exploratory tool for investigators and a decision-support tool for physicians and clinicians. Hopefully, the method can lead experts to make more effective decisions about localizing seizures, describing epileptic syndromes, and designing therapy for patients with epilepsy.

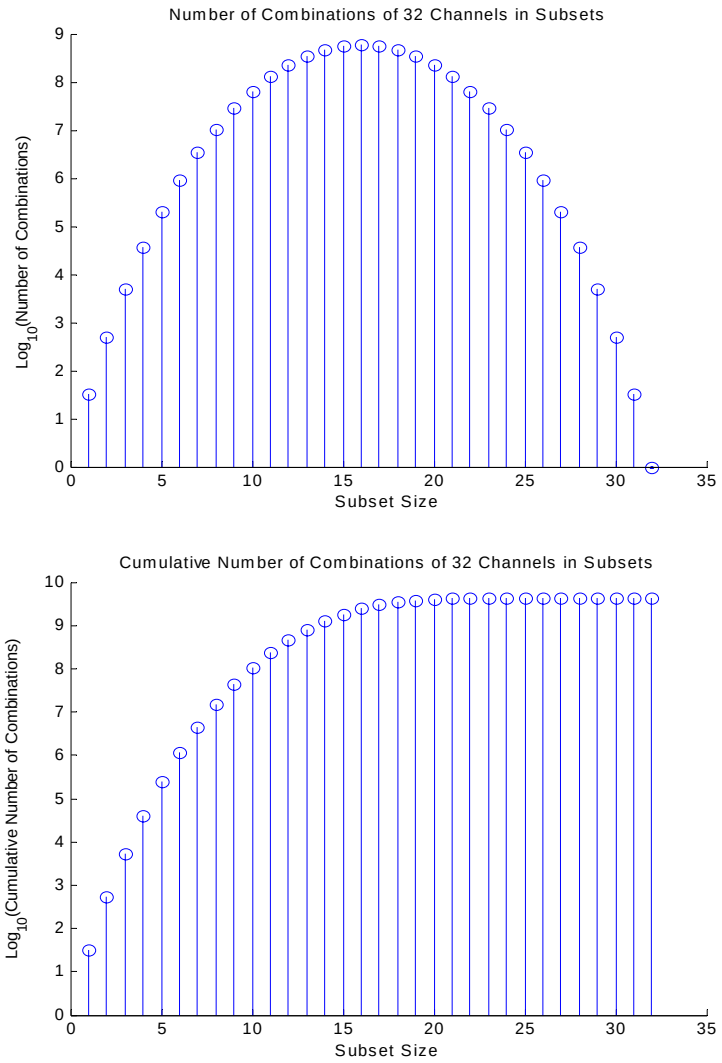


Figure 1.2-1 This figure depicts the value of using FIM to statistically cluster channels of IEEG with high incidences of epileptic events. As an example, the figure illustrates the single (*top*) and cumulative (*bottom*) number of combinations in a set of thirty-two channels. The FIM essentially filters through all possible combinations—ranging from tens to nearly billions, returning the most frequently occurring subset of channels.

1.3 Method Overview

Figure 1.3-1 illustrates the methodology for mapping a putative epileptic network in a brain with seizures according to frequently occurring concentrations of epileptic oscillations. The method incorporated two major stages: 1) the detection of pathological waveforms, and 2) the statistical clustering of the detected waveforms.

As illustrated in the below figure, raw IEEG data is fed into the algorithm for the binary detector, which ultimately returns the beginning and ending times and locations (electrodes) of a detected event. The detector includes modules for enhancement to highlight the oscillations that may be obscured within the typical activity of the IEEG data; extraction of a feature—computed off-line with an evolutionary algorithm—to transform the data into a quantifiable series that best distinguishes epileptic and non-epileptic interictal activity; and classification of the feature-series to determine the locations and time intervals containing an epileptic oscillation.

For each detected interictal event, a physics-based feature of interest, or a measurement (e.g., energy, entropy, duration), can be computed and compiled in a spatial-temporal matrix. It is hoped that the measure is an apt quantity to correlate the concentration of epileptic oscillations and location of epileptic networks involving seizures. In the measurement matrix, the columns and the rows represent the reference channel numbers in the raw IEEG data and the intervals of time (e.g., seconds, minutes) over which the measures are compiled, respectively. More specifically, every interval from the beginning until the end of the IEEG data record, a one-minute window slides over the detected events in the record and computes a statistic (e.g., the average energy,

total entropy) of the detected events per channel. For each interval and channel, a value is stored in the corresponding row and column of the measurement matrix.

Finally, the matrix of measurements is fused along the column direction per row according to a condition, such as "pick the top 10 percent of all energy measurements," to create a special set of channel numbers that over time are candidates for an epileptic network of dysfunctional brain. The time-varying set of channels from the previous stage, are the input to the mining algorithm, which discovers patterns involving select channels during the IEEG record. For this task APRIORI, a popular algorithm for mining data, is implemented with permission from Bodon [153]. Although APRIORI will be treated as a "black-box" in this research, the details of the operation of this module can be found in several sources [153, 155-157]. A brief summary of APRIORI is provided in Appendix E. In short, the mining algorithm outputs the most frequent patterns involving certain channels according to a thresholded minimum likelihood of occurrence. That is, APRIORI can be used to discover patterns that happen within the IEEG 35%, 80%, or even 100% of the time as well as any other specific rate of occurrence.

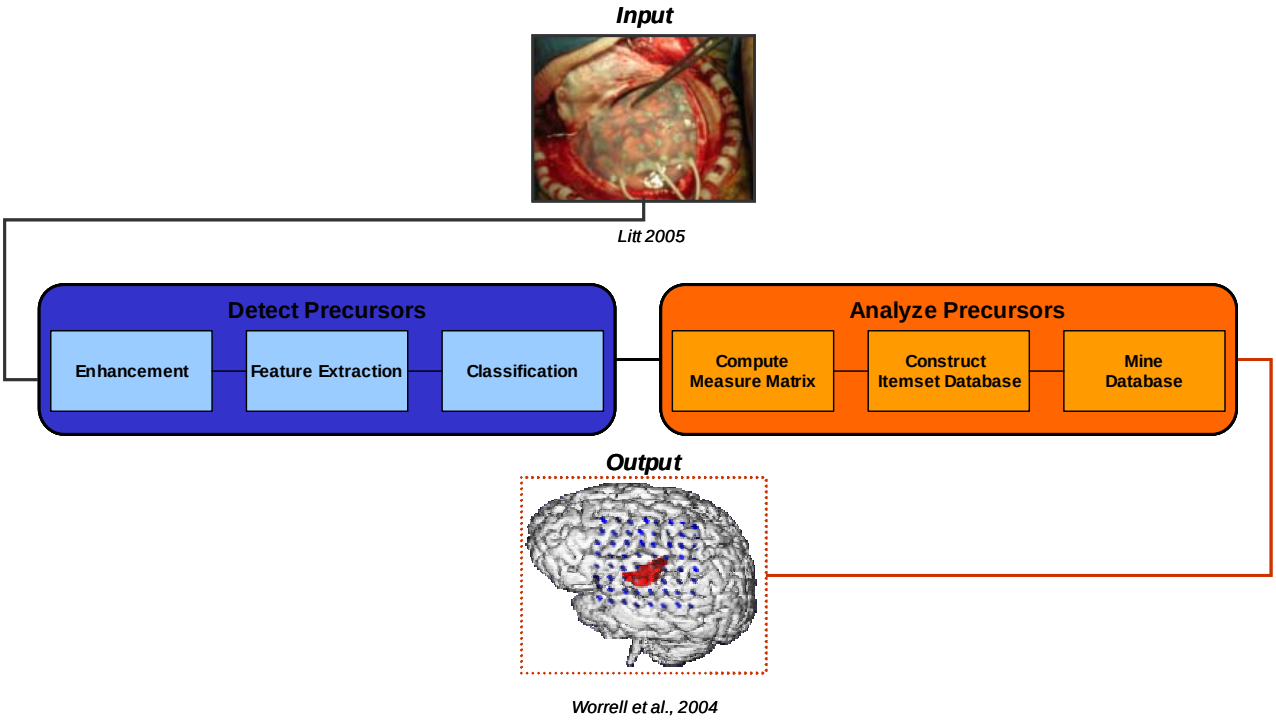


Figure 1.3-1 This figure illustrates the two-stage methodology for pinpointing considerably dysfunctional sites of brain based upon detecting concentrations of epileptic oscillations. Each of the two major stages, detection (*blue*) and analysis (*orange*) possesses three modules. In the first stage, pathological oscillations within intracranial EEG recordings (*input*) across several channels and intervals of time is processed by filtering, extracting feature values, and classifying the extracted values to yield the instances in time and location of an oscillation. For the second stage, measurements of all the detected abnormal activity are made and compiled across space and time that may be mined to reveal any frequently occurring patterns (if they exist) that can be mapped as the principal areas of dysfunctional brain (*output*). The final output serves to support the decisions made by a physician.

1.4 Solution Uniqueness

The epilepsy literature does not present an application that uses both FIM and an EA to localize epileptic networks. In fact, the same literature does not apply techniques in pattern recognition to the particular EHFO that this work examines. Moreover, the existing epilepsy literature does not reveal the application of PSO for optimally distinguishing between abnormal and normal events in multi-electrode IEEG, but does discuss two primary efforts to apply GP to biological signals and one effort in to apply FIM for discovering patterns in depth electrodes.

In 1997, Marchesi et al. propose a method using GP for the automatic detection of spike-and-slow-wave complexes that occurred during the discharges of absence (petit-mal) seizures. The method, which inputs a small set of classical (conventional) features, or c-features, achieved a sensitivity of 73%, specificity of 97%, and selectivity of 96% on a small database (80 lone events) of testing patterns. The following points distinguish their work from this research:

- The event of interest is essentially an ictal event, since the spike-and-slow-wave complexes transpire along with seizures, rather than an interictal event.
- The library of features is limited to only seven measures from only the time-domain (i.e., *mean absolute value*, *average value*, *number of zero-crossings*, *number of sign changes*, *upward slope*, *downward slope*, *duration*) with obvious correlations among the manually selected c-features, which introduces computational redundancy, rather than a larger set of features from different domains (Appendix B).

In 2005, Firpi, Goodman and Echauz develop algorithms for the detection and prediction of medial temporal lobe seizures relying on genetic programming. In each case, their algorithm outperforms a benchmark that uses the feature, *curve length*, which presupposes the physics of the events within the data. The following points distinguish their work from this research:

- The patients possess medial temporal lobe but not neocortical seizures.
- The event of interest is either an ictal event (seizure) or a preictal event (pre-seizure interval) of an arbitrary duration (~20 minutes).
- A subject-specific artificial feature is computed only from raw data but not using a library of features (Appendix B).

To date, the epilepsy literature reveals that only Bourien et al. apply frequent itemset mining to detectable interictal abnormal waveforms in the IEEG, although a few sources have used FIM for feature extraction and classification [151, 152, 159]. However, the following points distinguish their efforts from the current research:

- The patients possess medial temporal lobe but not neocortical seizures.
- The events of interest are spikes-and-sharp waves (>20 Hz) and "fast activity" (30-50 Hz), which are not necessarily EHFO, rather than HFEO's (>60 Hz).
- A single feature, *mean absolute value*, was assumed the best feature for a binary detector without comparison or proof against other features.
- The performance of the suggested detector is unreported or unverified.
- Only one hour of data over two hours removed from the time of ictal onset per seizure record is analyzed but not the interictal time immediately prior to onset.

1.5 Contributions

Overall, this research provides architecture for the automatic detection and clustering of epileptic oscillations in multi-electrode IEEG over long durations to ultimately benefit persons with epilepsy by guiding physicians and therapeutic techniques. The following items outline the contributions of this research:

- Evolutionary algorithms to select and fuse a set of classical (physics-based) features for a detector of epileptic oscillations.
- Evolutionary algorithms to create features derived directly from data (data-driven features) for a detector of epileptic oscillations.
- Demonstration that a detector using features created via evolutionary algorithms outperforms a detector relying on a standard feature according to sensitivity and specificity.
- Frequent itemset mining to cluster electrodes concentrated with a particular epileptic oscillation. Development of this framework may prove to localize the epileptogenic zone (EZ) in pre-surgical evaluation for the effective treatment of epilepsy and essentially mapping networks of epileptic neurons.

Furthermore, the following tools result from the characterization of a certain epileptic oscillation:

- Hidden Markov modeling to simulate interictal IEEG including the event.
- Brownian motion as an analogue to describe how degenerative neurons produce an action potential for the event during synaptic transmission.
- Frequent itemset mining to demonstrate that the best set of features to distinguish the event is subject-specific, thus likely not universal.

Chapter 2: Origin and History of the Problem

2.1 Epilepsy, Seizures, and Treatment

2.1.1 Epilepsy

Although the first documented case of epilepsy dates back to 2000 B.C., it is not until the 19th century that epilepsy is widely accepted as a disorder [2]. Today, the magnitude of epilepsy as a disorder is clearly understood. Epilepsy is second to only strokes as the most widespread neurological dysfunction, yet is still suffered by approximately 1% of the world population [1]. In fact, about 2,500,000 Americans—of whom 30% are people under the age of eighteen—possess some kind of epilepsy [3]. Epilepsy can be engendered by several causes and is characterized by the occurrence of repeated seizures, which can manifest in various pathological forms and as a range of physical symptoms. Identified causes of epilepsy include genetic disorders, cortical malformations (dysplasia), metabolic disturbances, traumatic injury, stroke, tumors, infection, alcohol abuse, and drug abuse [4].

2.1.2 Epileptic Seizures

Despite years of extensive investigation by doctors and researchers, little has been discovered on the underlying mechanisms of seizures, making epilepsy appear to be a huge puzzle requiring more pieces to aptly envision. What is known about epilepsy is that a seizure is a disturbance of the neuronal electrochemical activity such that a set of neurons suddenly produces a repetitive, synchronous discharge. Factors such as the manner of inception in the brain, the level of awareness during an onset, and the extent of

affliction categorize the seizures [5]. The category of a seizure influences its duration (usually a few seconds to several minutes) and the symptoms (e.g., involuntary movements, an altered state of awareness, convulsions, impaired consciousness, unusual or repetitive behaviors, abnormal sensations) endured by a patient [5]. The various classifications of epileptic seizures are illustrated in Figure 2.1-1.

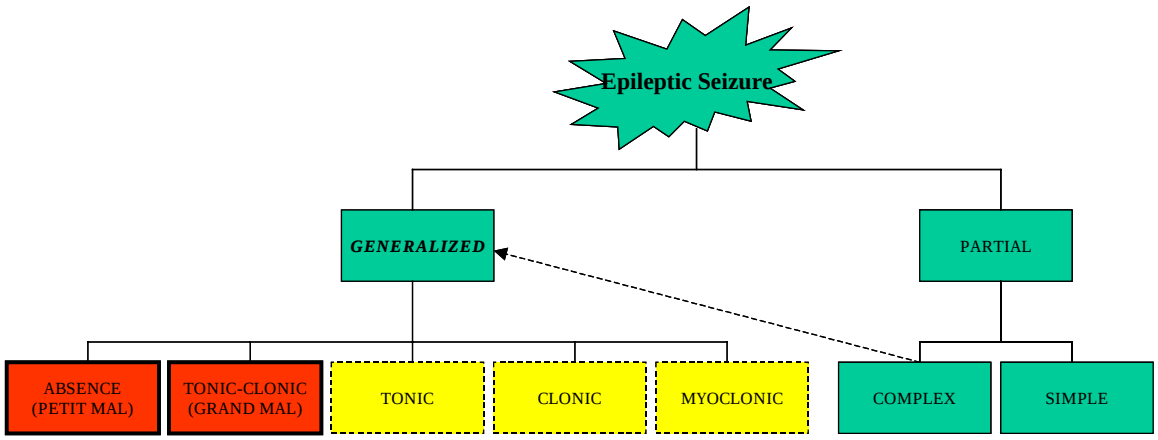


Figure 2.1-1 This figure illustrates the taxonomy of a seizure. The bold italic text denotes they type of seizures under study in this research. The red, bold-edged boxes and yellow, dashed-edged boxes indicate more and less common seizures, respectively. The dashed arrow between the partial complex seizure category and the generalized seizure class represents the possibility of a partial seizure degenerating to a generalized seizure.

As shown in the above figure, there exist two major types of seizures: partial and generalized. Partial seizures affect only a portion of the brain, originating from a focal region of brain. However, it is possible for a partial seizure to initiate in a focal region and disperse to other areas of tissue, making them resemble generalized seizures.

Patients with complex partial seizures experience no or impaired consciousness during a disturbance, while consciousness is preserved for patients with simple partial seizures.

In contrast, generalized seizures involve the entire brain, erupting throughout the cerebral cortex at one time and more often than not impairing consciousness [5, 6]. Generalized seizures can be further cataloged according to their clinical and electrical characteristics. Absence (petit mal) seizures are characterized by an immediate cessation of conscious activity often accompanied by a blank stare and eyelid fluttering without loss of posture. Tonic seizures are evident by an increase in tone, also referred to as tonic stiffening. Clonic seizures are distinguished by rhythmic muscle contractions or jerking. Myoclonic seizures are involuntary clonic movements affecting one body member or more. Finally, tonic-clonic (grand mal) seizures combine the effects of tonic and clonic seizures, in that order [9, 10].

2.1.3 Epilepsy Therapy

The complex taxonomy of seizures indicates the challenge in providing treatment for individuals with epilepsy. Although surgery is a highly successful treatment for epilepsy patients, only about eight percent of those persons—individuals with strictly partial seizures—undergo surgery. Furthermore, the lack of a reliable method for pinpointing the exact location(s) and spread of the pathology usually necessitates the removal of a significantly large volume of brain tissue. Additionally, surgery is a risky procedure, despite a highly selective patient population, boding a 10% chance of morbidity [11]. Another conventional recourse is anti-epileptic drugs, but risks adverse side effects in up to half of these patients [5, 38]. Unfortunately, the remaining 25% of

the persons with epilepsy cannot be treated with either medicinal or surgical care. The known limitations in providing therapy to this group include pathology that is poorly localized or exists in multiple focal regions, seizures that begin in vital irremovable brain matter, and the available drugs ineffectively quell seizures yet yield side effects. Figure 2.1-2 captures the aforementioned issues and evinces the need for new therapeutic solutions.

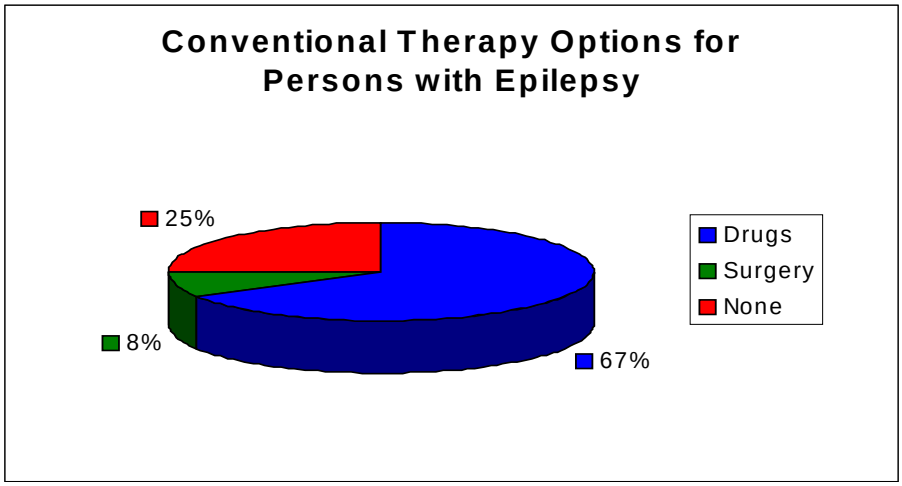


Figure 2.1-2 This figure illustrates the suitability of conventional treatment for individuals with epilepsy. Of considerable note is that the most dominant therapies fail to help approximately one-fourth of those who have epilepsy.

Attempting to create alternatives for previous treatment options, researchers of epilepsy therapy have concentrated on developing devices that supply local therapy in the brain to abort seizures. A number of these devices involve electrical stimulation [12-28], drug delivery [29, 30], focal cooling [39-42], or electric and magnetic fields [43-47]. Though

work in the area of epilepsy treatment has presented very promising techniques, the efforts still bear an essential difficulty in administering effective treatment: the ability to precisely localize the best brain region before application. Therefore, to resolve this fundamental concern and advance the field of epilepsy study, clinicians and researchers must thoroughly examine electrographic recordings, searching for relevant information that assists in localizing epileptic networks. Predominantly, this goal entails pinpointing the epileptogenic zone (EZ), or area of the brain that is compulsory in engendering epileptic seizures and overlaps with the seizure onset zone (SOZ), or area from which clinical seizures are actually generated.

2.2 The EEG and the IEEG

Recorded electrical activity of the brain, or the electroencephalogram (EEG), was first acquired from humans in 1929 [48], and since then the EEG has become the primary tool of epilepsy diagnosis and seizure analysis. The EEG voltage can be considered a random process, grossly equivalent to a superposition of the extra-cellular potentials of neurons near the EEG electrode, while the neuron potentials are sample functions (realizations) contributing to the EEG voltage. It is believed that in a brain without epilepsy, neurons are constantly interacting—conducting, transmitting, and receiving electrochemical signals—in a disorganized manner [7]. Furthermore, it is believed that when a seizure begins in an epileptic brain, neurons interact more frequently and in a synchronous fashion. In the latter case, as the number of interactions increase, an accumulation of neuronal responses culminates in an abnormal discharge, which neurologists regard as seizures [7]. Figure 2.2-1 illustrates this apparent synchrony.

In the below figure, which represents a segment of actual IEEG, the blue plot lasting about nine seconds prior to the time mark at 0 seconds is baseline (background) activity, and the red plot is ictal (seizure) activity. Notice that the background activity appears to be white noise while the ictal activity is somewhat sinusoidal—and organized in a sense, evolving in amplitude and frequency after a sudden electrographic change.

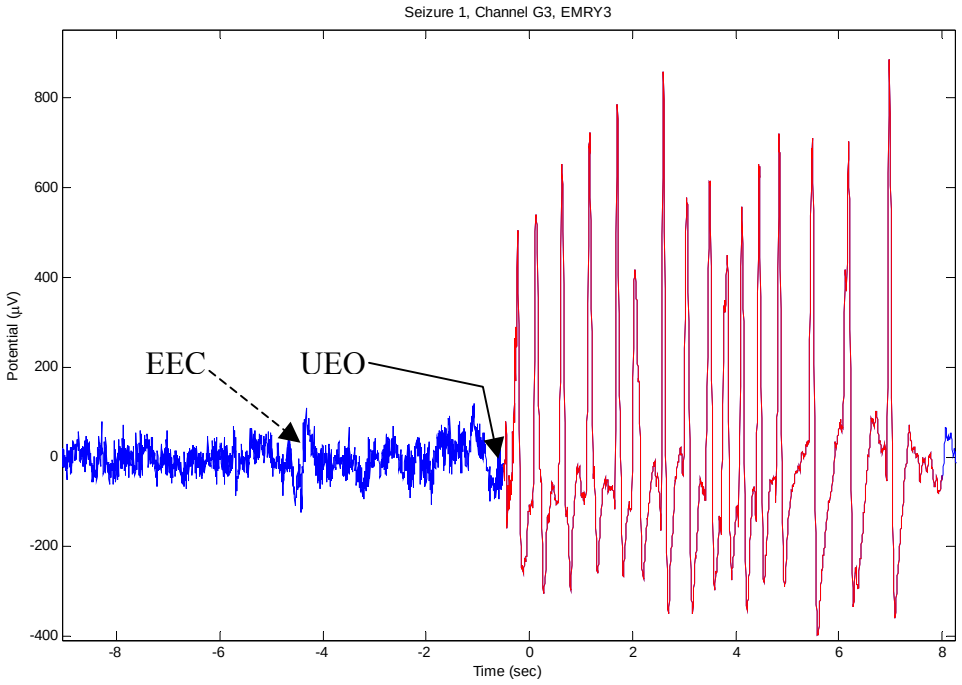


Figure 2.2-1 This figure illustrates the ideas of synchrony during a seizure (*red trace after solid arrow*) and asynchrony in the baseline activity (*blue trace before solid arrow*) that typically precedes a seizure. Often, epileptologists use the earliest electrographic change (EEC) and the unequivocal EEG onset (UEO), both distinguishing events from baseline activity, as key moments in time to study ictogenesis.

There are several ways in which the EEG is recorded, all differing primarily by the location of the measurement electrodes. The most common orientations of electrodes are surface placement and intracranial placement. Surface (or scalp) electrodes are affixed to the scalp of the epilepsy patient, recording electrical activity from the brain at a depth of about 1 cm below the surface of the brain. The attachment to the scalp makes surface electrodes a noninvasive recoding approach but also presents drawbacks that constrain their use for many EEG studies. First, the scalp recordings filter out much of the high-frequency signal content and distort amplitude measures because scalp EEG does not penetrate the skull, meninges, and subarachnoid space. Second, scalp EEG recordings are susceptible to motion, muscular, and recording artifacts. Third, and most importantly, scalp EEG cannot provide precise localization of primary pathological regions of the epileptic brain.

Intracranial electroencephalogram (IEEG) electrodes are surgically implanted within the brain tissue of epilepsy patients. From a research viewpoint, the depth to which the IEEG electrodes can be inserted depends upon the anatomical structure (e.g., hippocampus, amygdala, cingulate cortex) within the brain to be studied. However, ethical and safety considerations ultimately direct the use of the IEEG, reserving intracranial electrographic measurement for pre-surgical evaluation according to stringent standard protocols [49-52]. For instance, potential risks of damaging or infecting the brain curtail spatial sampling (number of electrode placements), and in turn the amount of information that can be gained from electrical activity in the brain. Nonetheless, the IEEG provides much better spatial resolution than surface electrodes, is nearly artifact free and has a high signal-to-noise ratio. Moreover, studying the IEEG has lead to the

discovery of electrographic waveforms that epileptologists speculate can localize or predict the onset of epileptic seizures.

2.3 Electrographic Waveforms

Since the ethical approval of using intracranial electrodes, clinicians have visually recognized interesting waveforms within the IEEG. Figure 2.3-1 contains representations of some of the events that have been monitored from the IEEG. While the majority of the observed signals seem rhythmic in shape and indicate normal activity of the brain, signatures with a relatively irregular morphology and epileptic-like incidence are also apparent [8].

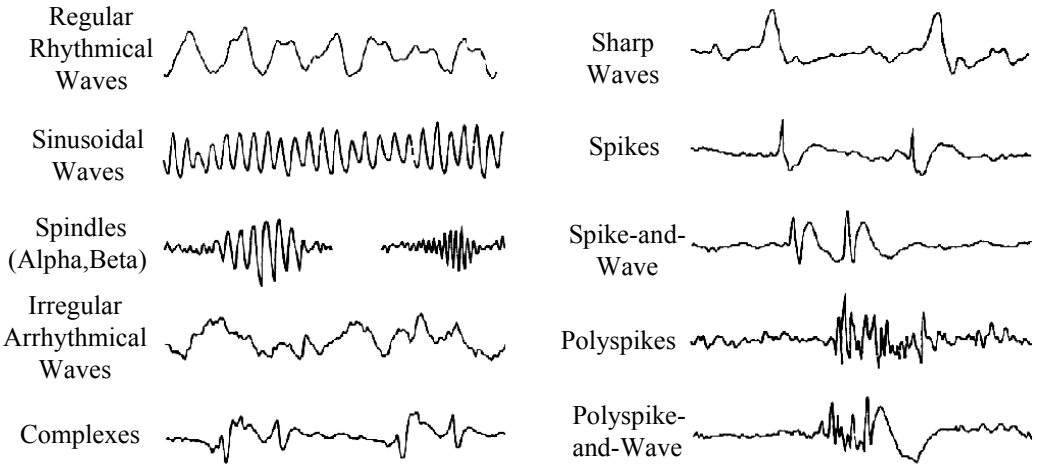


Figure 2.3-1 This figure displays a few of the signals observed in IEEG recordings. Some of the IEEG events are rhythmic while other signatures are arrhythmic and even seem "seizure-like" or pathological in character.

These irregular waveforms have been of great interest to neurologists who hypothesize that they are related to ictogenesis, localize seizures, and can forecast ictal onsets [53, 86, 95, 105-109, 120]. Although researchers have primarily considered sharp waves, spikes, and spike-and-waves for studying epileptic seizures, those investigations have achieved marginal success in automated detection of the abnormal waveforms [130-132, 137, 138, 141, 144, 145] and analysis of detections to understand epileptic mechanisms [113]. On the contrary, more recent findings suggest two types of abnormal waveforms, fast ripples and high frequency epileptiform oscillations, as substantial clues to aid diagnostic (and maybe even prognostic) tools for epilepsy research and treatment [53, 105-109, 136].

Like most abnormal waveforms in the IEEG, fast ripples and HFEO's may be characterized as a brief burst, lasting milliseconds to a few seconds, with low signal-to-noise ratio and electrographic activity that is similar in appearance to polyspikes and polyspike-and-wave signatures in the IEEG. More importantly, fast ripples and HFEO's may be key to spatially localizing the EZ and revealing the mechanisms of ictogenesis. Studying the 80-500 Hz frequency band from the hippocampal and entorhinal cortical areas of rodents and humans, Staba and collaborators suggest that fast ripples possess a strong association with regions of brain involved in ictogenesis [105-109]. Similarly, Worrell and researchers [53] find that high-frequency epileptiform oscillations appear to localize the seizure onset zone (SOZ), a sub-region of the EZ, in six patients with neocortical epilepsy. Thus, each irregular event appears to be pathological and possess great potential in advancing epilepsy research.

2.4 Detection of Abnormal Waveforms

This section provides context on the role of detecting abnormal electrographic waveforms that may track or map epileptic networks in epilepsy research. In addition, this section summarizes some previous research in detecting certain abnormal waveforms in the EEG or IEEG. This research is closely associated with research to detect or predict epileptic seizures. More details on the detection [54-74] and prediction [75-96] of seizures can be found in several sources.

Essentially, designing a method to detect and analyze abnormal waveforms conceptually parallels designing a method to detect seizures, where the event of interest is the primary disparity in the framework of each problem. That is, both problems are pattern detection problems but the abnormal waveform in the former problem is a relatively brief interictal or preictal event (e.g. HFEO's, fast ripples, spikes) and the abnormal waveform of the latter problem is an ictal event (a seizure). Initially, the epilepsy research community shared the notion that seizures are highly random phenomena without any prior indication of spatial or temporal occurrence. But over time, many research groups have observed clinical symptoms and quantitative measures of irregular electrographic activity that appear to foreshadow seizure onsets [97-129]. These findings have invoked interesting discussions and methods for seizure prediction and the predictability of epileptic seizures, intending to foster an effective proactive therapy for epileptic seizures. Thus, some laboratories have concentrated on the detection and analysis of abnormal interictal (or preictal) waveforms in IEEG to investigate if a clear connection exists between the occurrence of the irregular activity

and the onset of a seizure, hoping to design prognostic and diagnostic tools based upon such information.

To date, schemes to detect irregular waveforms have been focused on spikes, sharp waves, or combinations of spikes and sharp waves as an event of interest. In 1991, Wilson et al. develop a spike-detection process that yields performance measures of 74% sensitivity, 84% specificity, and 54% selectivity (positive predictive value) across a database of eight patients [130]. In 1996, Jones et al. outline a real-time system for automated detection of spikes and sharp waves, claiming 95% sensitivity and 72% selectivity [145]; however the computation of the system performance measures appears inconsistent with the complete statistics cited within the work. Also, the construction of a testing data set with sufficient epileptiform and non-epileptiform signal seems inadequate.

In 1998, Van Hoey et al. investigated the localization of epileptogenic brain (source analysis) using EEG spikes, which are detected with two features that are derived from periodic mathematical computations [131]. Visual inspection reveals that the features share good utility for spike detection; but verification on a sparse testing set, consisting of seven spikes in a 20-minute record for one patient during rest, leaves room for more extensive work. In the same year, Ko et al. presented an EEG spike detection algorithm using three features (i.e., peak angle, amplitude, velocity) as input to an artificial neural network (ANN) with radial basis functions in the hidden layer [132], while Tarassenko et al. use a time-domain and frequency-domain measures (e.g., slope, sharpness, autoregressive model coefficients) as inputs to a multi-layer perceptron (MLP)

neural network [133]. Despite an educated consideration of the features that enter each classifier, each method suffers very high false positive rates.

The following year, Zhang et al. proposes an approach for detecting epileptiform waves in EEG that incorporates the continuous wavelet transform (CWT) as a preprocessing stage, feature extraction using conventional spike measures, and a fuzzy neural network [144]. Compiling over 300 four-minute records from 17 patients, the authors state a sensitivity of 94% and selectivity of 88% [144]; however the approach combines multiple detections across channels and time samples into a single decision, masking the true values of the performance measures. Moreover, the method is not automated because visual inspection is needed to determine the appropriate scales for the CWT.

In the new millennium, most researchers remain with the old problem of spike detection. For instance, Calvagno et al. studied sub-band decomposition principles and the nonlinear energy measure; Pon et al. applied mathematical morphology and the wavelet transform; Ossadtchi et al. propose independent component analysis (ICA), the Recursively Applied and Projected MUltiple SIgnal Classification (RAP-MUSIC) algorithm, and nearest-neighbor clustering [146]; Van Hese et al. add features based on singular value decomposition (SVD) to earlier work on source analysis [141]; Hassanpour et al. study Fourier analysis in the time-frequency domain along with SVD and the nonlinear energy feature [138]; and Adjouadi et al. introduce the Walsh transform as a preprocessing module [137]. Although larger ranges of signal-processing techniques are explored, no convergence to a reliable prototype for spike detection is evident. In fact, many of the suggested methods for spike detection are not thoroughly tested (e.g.,

little to no cross-validation results, highly selective patient sets), making the approaches interesting but preliminary. The contemporary course of spike detection may aid in manually screening the EEG by a clinician or epileptologist but likely may not facilitate an apt analysis of the EEG to map epileptic networks with abnormal waveforms. This warrants a truly new direction in the detection of abnormal electrographic events that only a few groups have undertaken, namely the detection of epileptic high frequency oscillations.

In 2002, Staba and colleagues detect and analyze interictal epileptic oscillations between 80-500 Hz within the hippocampus and entorhinal cortex of humans with focal epilepsy. Though only reporting a sensitivity of 84% for their detector, their work demonstrates an association between the oscillations and epileptogenic regions and a statistical characterization of the signatures [136]. In 2003, Niederhauser et al. detect synchronous bursts of activity (20-40 Hz) in humans with focal epilepsy in the medial temporal lobe. While reliable detection is achieved in five patients with unilateral seizure onsets, the method does not recognize the event of interest in five other patients with more generalized onsets; and the events are not noticeable beyond two minutes prior to electrical onset time [139]. In 2004, Fischer and Tetzlaff propose a universal pattern detection procedure for oscillations based upon a template of 512 patterns drawn from one week of EEG recordings. Their approach requires an obviously impractical phase of screening days of data to construct an essentially exhaustive set of EEG patterns, some of which may last milliseconds in duration. In 2004 and 2005, Smart et al. investigate epileptiform oscillations between 60-100 Hz before generalized seizure onsets in humans. The group reports that the oscillations localize regions and moments before the earliest

electrographic change in depth recordings for 77% and 61% of the observed ictal onsets, respectively; however, the recommended detector is only 68% selective [142, 143].

In summary, there is still much room to advance the "state of the art" in the detection of oscillations, primarily by discovering better features for extraction and input to a classifier. This research presents evolutionary algorithms as a plausible solution to the abovementioned problem. The details of two evolutionary algorithms may be found in Appendix C and Appendix D.

Chapter 3: Methodology

This chapter discusses in detail a general methodology for the detection of abnormal electrographic waveforms that is applied to a specific type of waveform: an epileptic high frequency oscillation termed a high frequency epileptiform oscillation (HFEO). In short, the detector included four major modules: 1) a stage for signal enhancement, typically called pre-processing, to accentuate potential EHFO's within the time-series that was input to the detector; 2) a stage to extract feature values from the input; 3) a stage to classify the feature values as points in a probability distribution for either a class of an EHFO or a class non-EHFO (background activity); and 4) a stage to simply note and catalogue the beginning and ending times of a detected event. Furthermore, this chapter includes the framework of a proven frequent itemset mining algorithm, APRIORI, to statistically cluster concentrations of epileptic activity having processed long-duration, multi-channel interictal records containing events of interest, in this case HFEO's, detected with the established methodology for robust detection.

All work toward designing, testing and validating the automated detector of HFEO's is done using MATLAB version 7.0.4.365 (R14) software on a Dell Dimension 8300 desktop that is equipped with an Intel Pentium 4 processor (3.2 GHz) and 1 GB of RAM, while the algorithm, APRIORI, is freeware that executes in C.

3.1 Data Collection

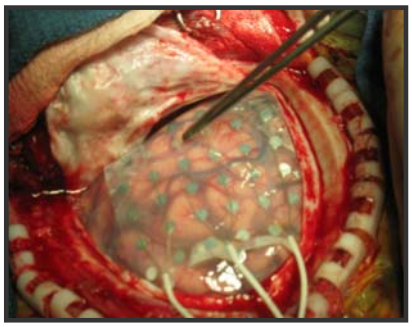
Seven patients who underwent long-term continuous video and intracranial EEG monitoring at Emory University and University of Pennsylvania as part of their

evaluation for epilepsy surgery were studied. Each patient gave their informed consent for participation in this investigation under the approval of the respective Internal Review Boards of the universities. The IEEG data were de-identified before analyses.

The data for each patient were collected using a digital, 64-channel, 12-bit, Nicolet BMS-5000 (Nicolet Biomedical Inc., Madison, WI) epilepsy monitoring system. Each Emory University (i.e., Emory Hospital) and the University of Pennsylvania (i.e., the Children's Hospital of Philadelphia) housed an epilepsy-monitoring unit (EMU) to collect hours of data while a clinician monitored patients. Referentially recorded EEG was band-passed filtered from 0.1-100 Hz and digitized at 200 Hz before archiving to CD-ROM for later processing. Bipolar electrode montages were used to reduce common mode artifact, and a digital 60 Hz notch filter was implemented to eliminate line noise.

Marking the times that seizures and epileptic oscillations preceding seizures happened was important for reviewing the results of the methodology to map abnormal activity, compute artificial features using the evolutionary algorithms, and validate the binary detector. Seizure onsets for the entire archive of data were marked by two board-certified epileptologists (G.A.W. and B.L.). The doctors determined the onset times for each seizure by looking backwards in the record for the times of the EEC and UEO. The doctors defined the seizure onset zone (SOZ) as the set of electrodes proximal to the electrode exhibiting the EEC. On the other hand, registering all interictal HFEO's across sixty-four channels of EEG in several patients was an impractical task, especially since the duration of the oscillations was on the order of milliseconds. Therefore, two epileptologists (G.A.W. and E.M.) annotated a subset of records with HFEO's with short duration (at most 3.5 minutes) per patient.

Three subsets of data were apportioned from the above collection of data. A set of *training data* that supported the creation of artificial features via evolutionary algorithms and the classification of new values from a reference of values. The training data per subject consisted of randomly selected clippings from continuous EEG to compile a single record with at least 40 epileptic oscillations complemented by one minute of continuous background activity. One set of *testing data* aided in validating that the method could work well given data that are not known a priori. This set consisted of randomly constructed records of IEEG containing epileptic oscillations surrounded by (before and after) one second of electrographic background in a single channel. Finally, a second set of testing data with each record spanning several minutes to a couple of hours before the onset of a seizure over multiple channels of IEEG provided a means to study the ability of the method to pinpoint specific regions of dysfunctional brain that could lead to a better understanding of ictogenesis, electrode placement, and treating epilepsy based on statistics of high frequency epileptiform oscillations concentrated in the regions.



Litt, B.

Figure 3.1-1 This figure depicts the collection of intracranial EEG recordings from patients with neocortical epilepsy. Neocortical epilepsy affects the neocortex, which transmits and receives electrochemical signals from neurons at the surface tissue of brain.

3.2 Signal Enhancement

Worrell et al. remarked that the frequency band of an HFEO was between 60 Hz and 100 Hz in adult patients with neocortical epilepsy [53]. In addition, Marsh has noticed that similar epileptic waveforms in pediatric patients oscillate within 50 to 85 Hz [150]. Thus band-pass filtering the IEEG to improve the signal-to-noise ratio between the activity of an epileptic slow ripple and the activity of typical electrographic background was employed. The best options for simple filter design included a Chebychev filter with a small pass-band ripple and narrow transition bands (sharp cut-off frequencies) and a Butterworth filter with no ripple in the pass-band but relatively broad transition bands (blunt cut-off frequencies). Because the best contrast between background activity and epileptic activity implied that some low frequency content was needed and a filter with constant magnitude maintained the amplitude of each oscillation for stable detection, the Butterworth filter with tolerable transition bands and even gain fit the needs of the detector. Phase distortion—which can introduce time-delay—from the filter was eliminated by filtering the data, filtering the reverse of the filtered data, then reversing and rescaling the output of the last operation of filtering.

Another consideration for enhancement was that band-pass filtering would augment some broadband, physiological clatter that could be mistaken for a pathological fluctuation. With the morphology of most HFEO's resembling an approximate sinusoid, time-differentiation, also referred to as whitening or equalization, of the IEEG was anticipated to aid in discerning an HFEO from background activity. Of course, whitening could also enlarge ordinary, undesired oscillations or broadband spiking transients within the IEEG. Therefore, band-pass filtering and whitening were cascaded to exploit the

benefits of each technique for enhancement. Figure 3.2-1 shows that this enhancement produced by whitening then band-pass filtering successfully highlighted EHFO's in the original time-series for a single channel of IEEG data.

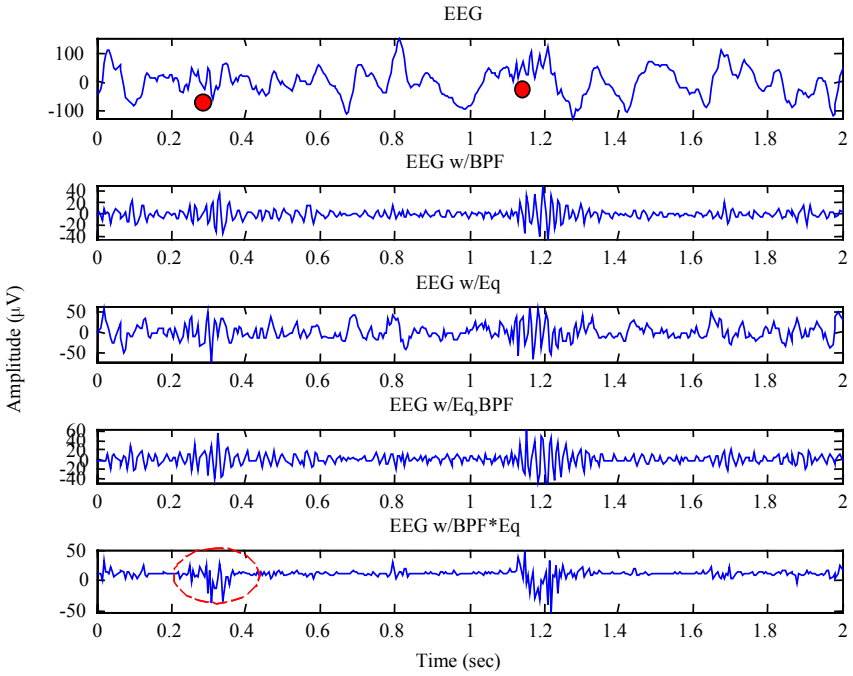


Figure 3.2-1 This figure shows the effectiveness of the techniques for emphasizing the activity of an EHFO within the EEG (*top*) of a selected patient. The 2nd and 3rd panels plot the band-pass filtered and whitened EEG, respectively. The final two plots accordingly represent the final and an alternative means for enhancement of the IEEG. A dot indicates an EHFO.

The above figure illustrates in a single observation that band-pass filtering of a whitened series of EEG was sufficient for signal enhancement, a result that held strongly for several other cases, when reviewing the contrast between the expert markings and

background. Interestingly, scaling the product of the whitened signal and the filtered signal appeared to greatly improve the signal-to-noise ratio of the original EEG as demonstrated by juxtaposing the time interval between 200 and 400 milliseconds in all time-series of figure 3.2-1. More specifically, the first event (left-most dot from the top plot) of the EEG stood out the most using the enhancement of the last panel. Nonetheless, band-pass filtering a whitened time-series was preferred over scaling a band-pass filtered signal multiplied by a whitened signal because the latter approach for signal enhancement modulates the spectrum of the original time-series, thereby obscuring the interpretation of features in the frequency-domain.

3.3 Feature Extraction

Figure 3.3-1 illustrates the operation of extracting a feature, which was typically performed in a causal moving window fashion. A sliding window of length W defined the number of values—the current and previous W-1 points in time—from a discrete-time input, $x[n]$, to evaluate with a function, Φ, that describes the feature to extract. As the window was shifted D points in time, the feature processes a new block of data. If the length of $x[n]$ was P points, then feature extraction yielded $\mathrm{K} = 1 + \left\lfloor \frac{P-W}{D} \right\rfloor$ points in a vector with $R = P - (D \cdot (K-1) + W)$ points remaining unprocessed from the original discrete time-series, presuming that $1 \le D \le W \le P$.

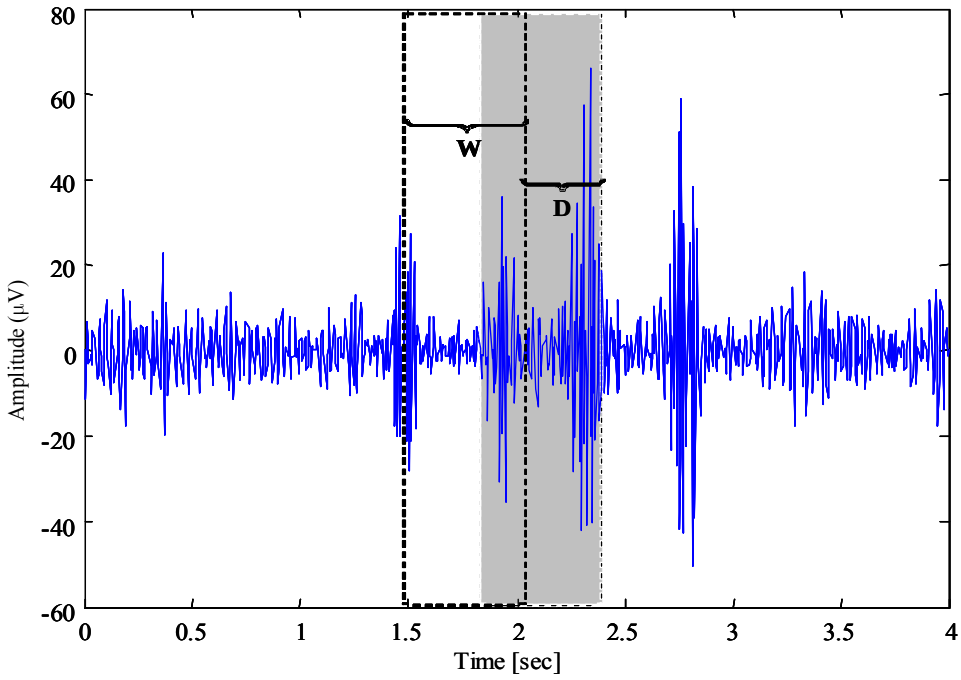


Figure 3.3-1 An illustration of the procedure for feature-extraction, where the equation for a feature operates on the EEG values within a window of length *W* (*shaded rectangle*) as the window is displaced (*dotted rectangle*) in intervals of length *D*.

Typically in literature for detecting epileptic oscillations, the length of the sliding window was arbitrarily set given knowledge about the duration of the event of interest and presumed stationarity in the original time-series. This arbitrary setting was surprising considering the importance of the correct window in extracting a feature. In this work, a statistical approach was employed to estimate the proper value for *W*. That was, the expected value of the duration of an epileptic oscillation per subject was estimated by bootstrapping the average of a sample of measured durations, and *W* equaled a value below—but close in value to—the mean. Figure 3.3-2 and Table 3.3-1 represent the results of such an analysis.

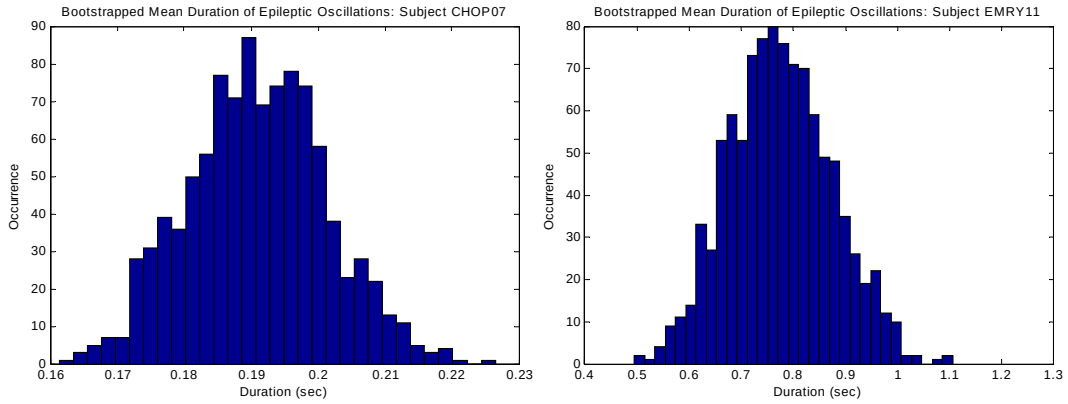


Figure 3.3-2 This figure illustrates the use of bootstrapping to compute a confidence interval for the expected duration of a specific epileptic oscillation in two subjects. This analysis was used to guide the selection of a value for the sliding window in extracting and creating features for a binary detector for each subject. Juxtaposing the left and right panels of the figure suggests that the sliding window is subject-specific.

For instance, sliding windows of 160 and 500 milliseconds—the lower bounds of the histograms—for the left and right panels of Figure 3.3-3 were chosen to extract features for events expected to last about 190 and 760 milliseconds, respectively. The minimum value of a histogram was taken as an appropriate size for the window in understanding that a window larger than the true duration of an oscillation would always blur the contrast between events and non-events in the feature-series and quash a major purpose of extracting a feature. Thus, choosing a window much smaller than the mean duration of an event was a reasonable precaution.

The histograms per subject were computed by bootstrapping (or sampling with replacement) a sample of durations (N≥40). From the bootstrapped sample, a confidence interval (CI) in the expected value of an epileptic oscillation with 5% error was

calculated. Generally speaking, bootstrapping was a robust technique to estimate the parameters (statistics) of a sample when the distribution of the population from which the sample was drawn is unknown [178, 179]. An alternative estimate for the 95% CI used the original sample of durations [177-179] as described in the following equation:

$$95\%\,CI = \bar{x} \pm 2 \cdot \frac{s}{\sqrt{n}} \tag{3.3-1}$$

where $\bar{x}$ was the mean, s was the standard deviation, and n was the size of the sample. Table 3.3-1 contains values of the CI using bootstrapping and Equation 3.3-1.

Table 3.3-1 This table holds estimates for the 95% CI for the mean duration of an epileptic oscillation for each patient using bootstrapping (CI^1) and Equation 3.3-1 (CI^2) as well as an appropriate window (W) for extracting a subject-specific feature. The units of the values are seconds.

Subject	W	95% CI^1	95% CI^2
C07	0.1576	0.1914 ± 0.0201	0.1913 ± 0.0204
E02	0.4135	0.4745 ± 0.0355	0.4746 ± 0.0357
E03	0.1602	0.1970 ± 0.0246	0.1971 ± 0.0249
E05	0.2050	0.3425 ± 0.1086	0.3427 ± 0.1104
E07	0.1797	0.2477 ± 0.0424	0.2478 ± 0.0431
E09	0.3048	0.5383 ± 0.1058	0.5377 ± 0.1067
E11	0.4229	0.7664 ± 0.1983	0.7680 ± 0.2020

The displacement, D, which generally was a more subjective parameter of feature extraction than W, controlled the compression of the data by (and the resolution of) the feature. Setting this parameter was important for time and memory constraints of processing, but did not necessitate a rigorous analysis. Strictly speaking, the major guiding consideration in setting the shift of the window was that $1 \le D \le W$. This interval for D allowed the sliding window to move (lower bound of 1) without ignoring information from $x[n]$ (upper bound of W). In this research, the displacement equaled one-fourth of the subject-specific window size, or $D = \left\lfloor \frac{W}{4} \right\rfloor$, for extraction.

As an aside, when $D = 1$, canonical feature extraction took a considerable amount of time to compute a series of feature values, or *feature series*, since the original time-series was minimally compressed to $\mathrm{K} = 2 + P - W$ points. On the other hand, in the case where the k^{th} point in a feature series, $y[k]$, was computed in the following manner:

$$y[k] = \frac{1}{W} \sum_{n=(k-1)}^{n=(W-1)+(k-1)} \phi\left(x[n]\right) \tag{3.3-2}$$

where ϕ was a *memoryless* function, or system in which the output of ϕ depended only on the original time-series, $x[n]$, at the same point in time, n, the feature extraction was *malleable* (or the feature being extracted is a *malleable feature*), meaning that feature extraction was achieved by filtering $\phi\left(x[n]\right)$ with a *finite impulse response* (FIR) W-point averaging filter. For instance, the malleable feature *mean energy*, produced a feature series with the following equation:

$$y[k] = \frac{1}{W} \sum_{n=(k-1)}^{n=(W-1)+(k-1)} (x[n])^2 \tag{3.3-3}$$

where, $\phi(x) = (x[n])^2$, or the square of the corresponding values of the input series. A smoother series, $y[n]$, similar to the series in Equation 3.3-3 can be obtained by convolving $v[n] = (x[n])^2$ with the following difference equation:

$$h[n] = \frac{1}{W} \sum_{i=0}^{i=(W-1)} \delta[n-i] \tag{3.3-4}$$

The convolution $y[n] = v[n] * h[n]$, however, resulted in *W-1* additional points at the end of $y[n]$ that was truncated because the sliding window "slid off" the data (which would have been the $R = P - (D \cdot (K-1) + W)$ points in canonical extraction) and *W-1* transitional points at the beginning of $y[n]$ that were considered a delay (as with canonical extraction) because the computation involved fewer than W points as the sliding window "slid onto" the data. In addition, non-causal implementation of the causal FIR filter in Equation 3.3-4 was used to remove the phase-distortion of the filter that introduced undesired time-delay in computing the feature series. That was, $v[n]$ was filtered in forward and in reverse using $h[n]$ to cancel phase-delay. It must be noted that the time-reversal of $v[n]$ ultimately made the filtering non-causal—although $h[n]$ was still causal, which would have been a practical limitation if $x[n]$ could not have been buffered.

In summary, a malleable feature provided a computationally inexpensive alternative to executing feature extraction that was implemented simply through filtering.

But a trade-off of using the filtering technique to extract a feature was the added burden of memory since the data was no longer compressed to at most $2+\left\lfloor \frac{P-W}{D} \right\rfloor$ points. Nonetheless, malleable features were advantageous and can be contrived in some applications of evolutionary algorithms. Figure 3.3-3 shows the results of extracting the *mean energy* feature from an enhanced EEG time-series with epileptic oscillations. In the 3rd panel of Figure 3.3-3, the feature series was computed according to Equation 3.3-2. In the 4th panel of Figure 3.3-3, the smooth feature series with zero delay was computed using the non-causal implementation of the causal FIR filter in Equation 3.3-4.

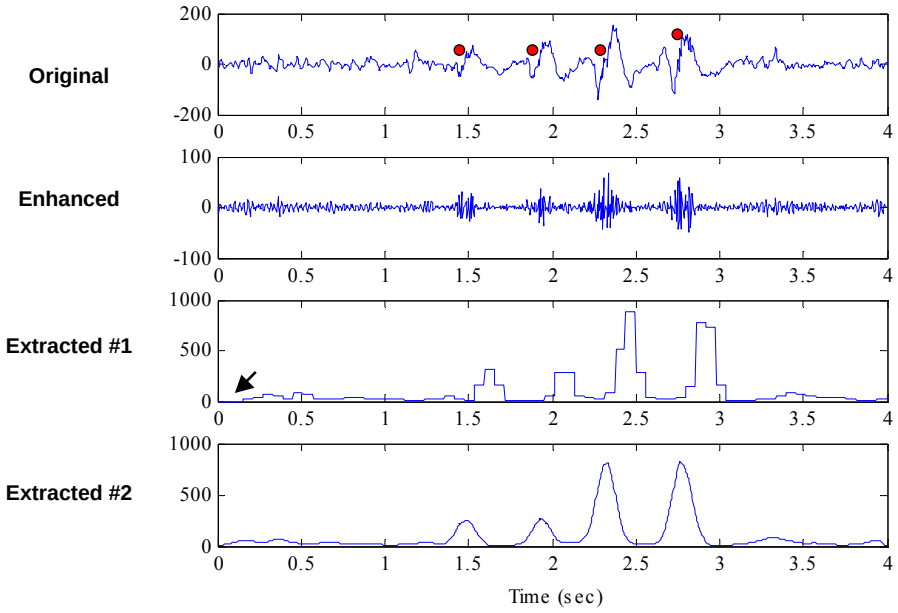


Figure 3.3-3 Feature extraction results for a four-second interval of EEG (*top panel*) that was enhanced (*second panel*) before computing the compressed series (*third panel*) and uncompressed series (*bottom panel*) using the feature *mean energy*. The red dots signify epileptic oscillations.

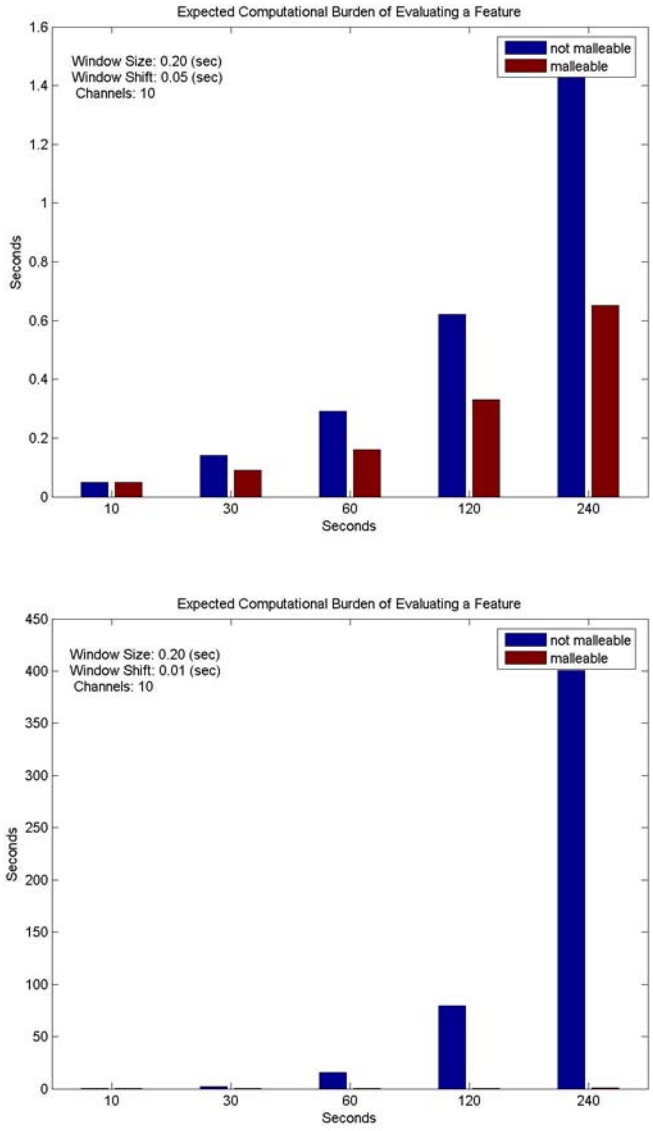


Figure 3.3-4 This figure illustrates the computational burden of extracting the feature mean energy by a malleable (*red*) and non-malleable (*blue*) implementation. The feature-vector (feature-series) is computed with a 200-millisecond-sliding window (W) and two different values for the displacement of the sliding window: $D = W/4$ (*top panel*) and $D = W/20$ (*bottom panel*). For either value of D, the malleable implementation is more efficient, especially in the second case where D is very small and the cost of classical (non-malleable) feature extraction increases extremely with the size of the data.

3.4 Feature Selection

Appendix B lists a set of features that were considered for feature extraction. These specific features were chosen because they were relatively simple to compute, computationally inexpensive to evaluate, and quantified some information that discriminated epileptic oscillations from non-epileptic oscillations. Next, feature selection was implemented to determine only those features in Appendix B with the best measurable information to detect epileptic oscillations.

The main purpose of feature selection was to choose the smallest subset of features to be used in classification while demonstrating an acceptable performance in classification [62, 184]. In Appendix F, a few approaches to feature selection were considered, including the application of a genetic program, because the epilepsy literature on feature extraction to detect epileptic oscillations does not present a benchmark. The GP proved to be a good technique for choosing (and combining) features.

3.5 Feature Creation

Feature extraction was considered the most important module of the detector. The effectiveness of the pattern classification in each subject relied on the use of an evolutionary algorithm to compute a high quality feature, meaning that the feature adequately quantified the difference between epileptic high frequency oscillations and background activity.

Each the genetic programming (GP) algorithm and the particle swarm optimization (PSO) algorithm described in Appendix C and Appendix D, respectively, was used to craft a highly discriminatory, artificial feature from the same training data,

which were enhanced EEG time-series obtained from seven subjects with neocortical epilepsy and interictal epileptic oscillations. Given the enhanced time-series for each patient, the GP operated on a matrix of values computed from either the classical features listed in Appendix B and the protected functions listed in Appendix A, so called GP-CF, or on a convolution matrix with five embedded delays, so called GP-DD, with parameters set to 35 generations, 350 individuals. Meanwhile, the PSO conveniently incorporated a neural network (NN) with two hidden layers (6x7 and 8x7 matrices of weights) and one output layer of sigmoid activation functions, yielding a PSO-NN algorithm. The implementation of GP-DD or PSO-NN provided a means to create a malleable feature for the detector.

Equation 3.5-1 was the chosen objective function, or *fitness*, that each evolutionary algorithm optimized to return a quality feature for extraction.

$$fitness = (1 - overlap) \cdot (K\,factor) \tag{3.5-1}$$

$$K\,factor = \frac{\left|\mu_1 - \mu_0\right|}{\sqrt{(\sigma_1^2 + \sigma_0^2)/2}} \tag{3.5-2}$$

$$overlap \;=\; \int \min(\;\; p(\mathbf{y} \mid \mathbf{c} = 0),\; p(\mathbf{y} \mid \mathbf{c} = 1)\;\;)\; dy \tag{3.5-3}$$

Equation 3.5-1 combined values from two metrics: the *k-factor*, which was a measure of separation between the prior probability distributions for the two classes of this problem, or essentially the Fisher's discriminant ratio (*FDR*) in one dimension; and the *overlap* between the prior probability distributions. In the equation for the *k-factor*, Equation 3.5-2, the mean (μ_c) and standard deviation (σ_c) of the probability distribution function

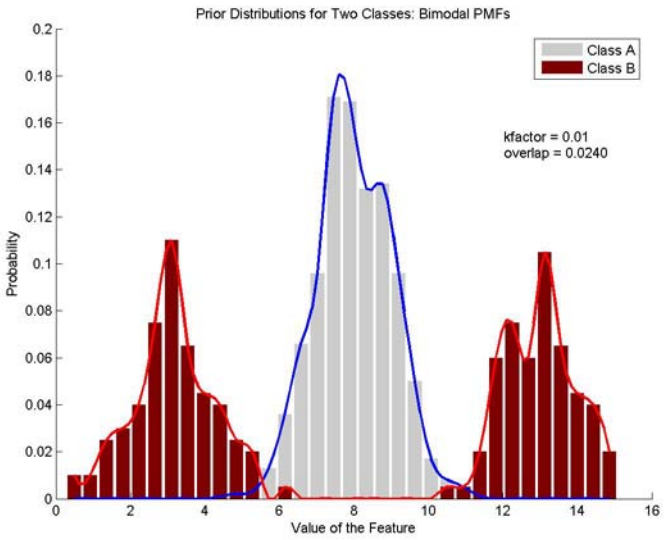
(PDF) for a group, *c*, with a label equal to 0 or 1 denoted the location and, spread of that group, respectively. The values of the *k-factor* range between 0 and infinity such that a high value represented great separation of classes. However, this metric alone was misguiding in cases where a multi-modal probability distribution for either group existed. In the equation for the *overlap*, Equation 3.5-3, the term $p(\mathbf{y} \mid \mathbf{c} = c)$ was the PDF of the feature **y** given that **y** belongs to a particular class *c*. The values of the *overlap* range between zero and one such that a low value represented good separation of classes. Thus, the fitness combined both metrics through multiplication to exploit the individual advantages of each measure.

There were two important points to make about the fitness in Equation 3.5-1. Firstly, conventional objective functions for evolutionary algorithms were classifier-dependent, or rely on the choice of a classifier through which the feature was passed to calculate a metric such as accuracy (or error) as the fitness, taking more time iteratively to create a feature that may not work well for any other classifier. On the other hand, Equation 3.5-1 describes a classifier-independent objective function that took less computational time to evaluate and was useful with any type of classifier. Secondly, the product of the overlap and k-factor prevented solutions with multimodal prior distributions exemplified in Figure 3.5-1a, although such a feature, Φ_0, was technically good. The cause of the omission the features in Figure 3.5-1a was the common mean feature value of 8 for the distributions of each class, which would result in a fitness of zero due to a k-factor of approximately zero. But, the evolutionary algorithm could have created an equivalent feature, Φ_e, represented in Figure 3.5-1b, where Equation 3.5-4

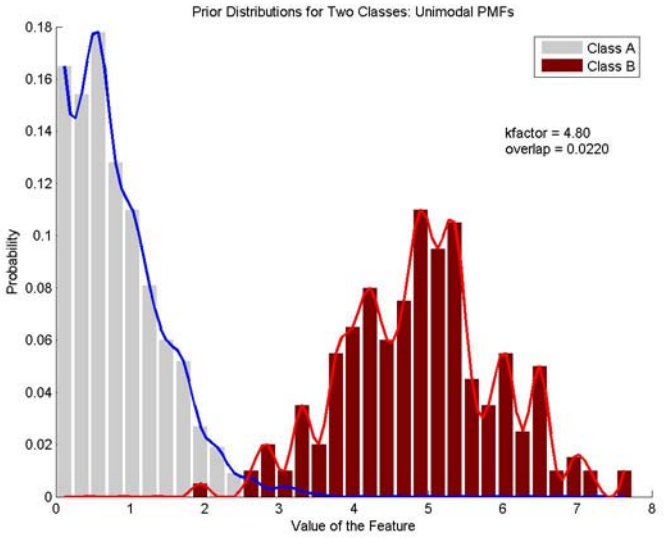
expressed the transformation of the values for Φ_0, v_0, that are centered at μ_0 to the values for Φ_e, v_e.

$$v_e = |v_0 - \mu_0| \tag{3.5-4}$$

Regardless of whether a classical feature (c-feature), such as *mean energy* or *kurtosis*, or an artificial feature (a-feature) derived via an evolutionary algorithm, such as GP, was considered for the detector of epileptic oscillations, the quality of the feature required an assessment. By computing a probability mass function (PMF) for a certain feature—or feature-space from which the mass function could be projected—and anticipating the performance in classification, the quality of the feature was readily determined. Figures 3.5-2, 3.5-3, and 3.5-5 portray the utility of an artificial feature against a classical feature in terms of estimated probability mass functions, orientation in the feature-space, and projected performance in classification respectively, for a single subject.



(a)



(b)

Figure 3.5-1 This figure illustrates the type of feature, Φ_0, that the evolutionary algorithms will not create *(a)* due to the chosen fitness—the product of the k-factor and the overlap between the prior probability distributions of the two classes in the detection problem—and a feature, Φ_e, that the algorithms could create *(b)* to be equivalent to Φ_0.

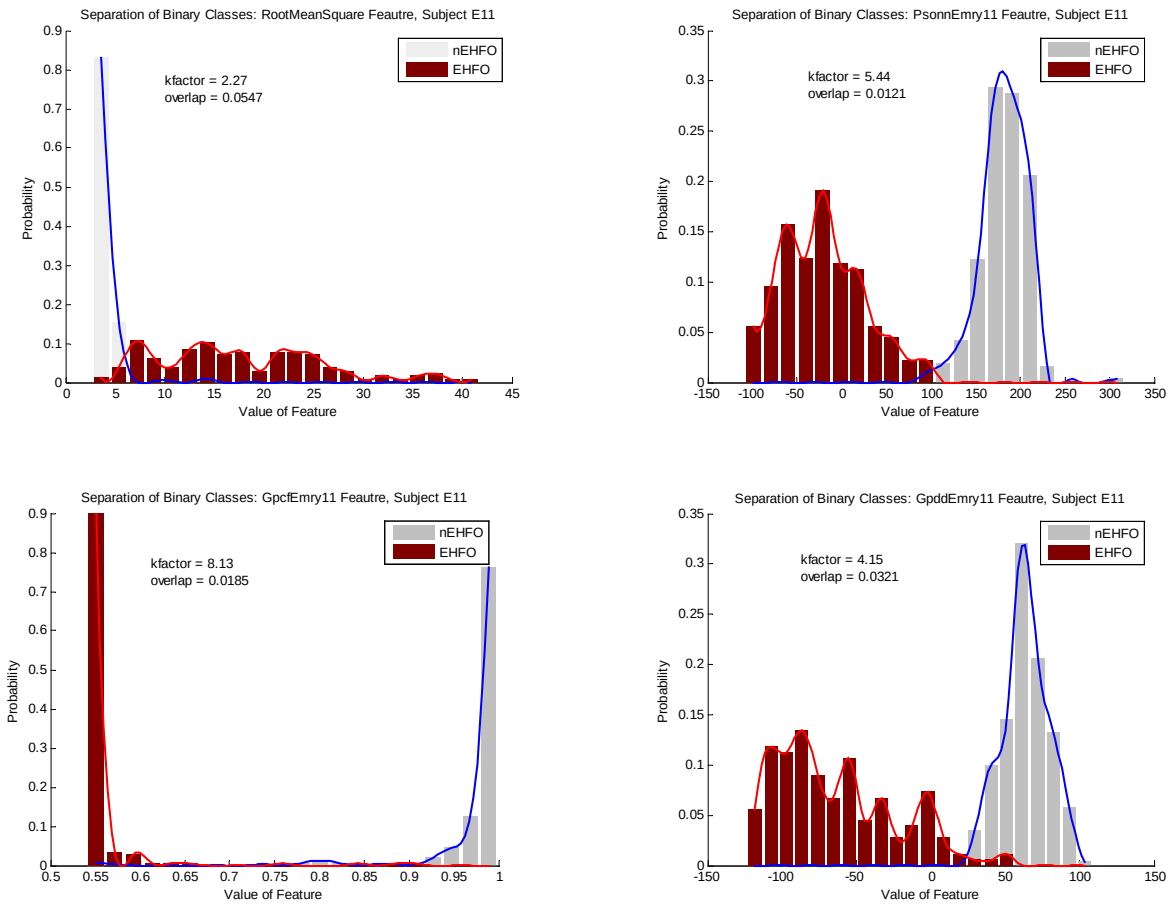


Figure 3.5-2 This figure illustrates for a subject the class-separation achieved by the presently accepted classical feature (*top left)*, genetically programmed artificial features (*bottom right, bottom left*), and a swarmed neural network artificial feature (*top right*) for detecting epileptic high frequency oscillations as represented by probability distributions. In this example, an artificial feature via GP-CF appears to exhibit the best separation as evinced visually and quantitatively with the metric k-factor, and all artificial features proved better than the RMS.

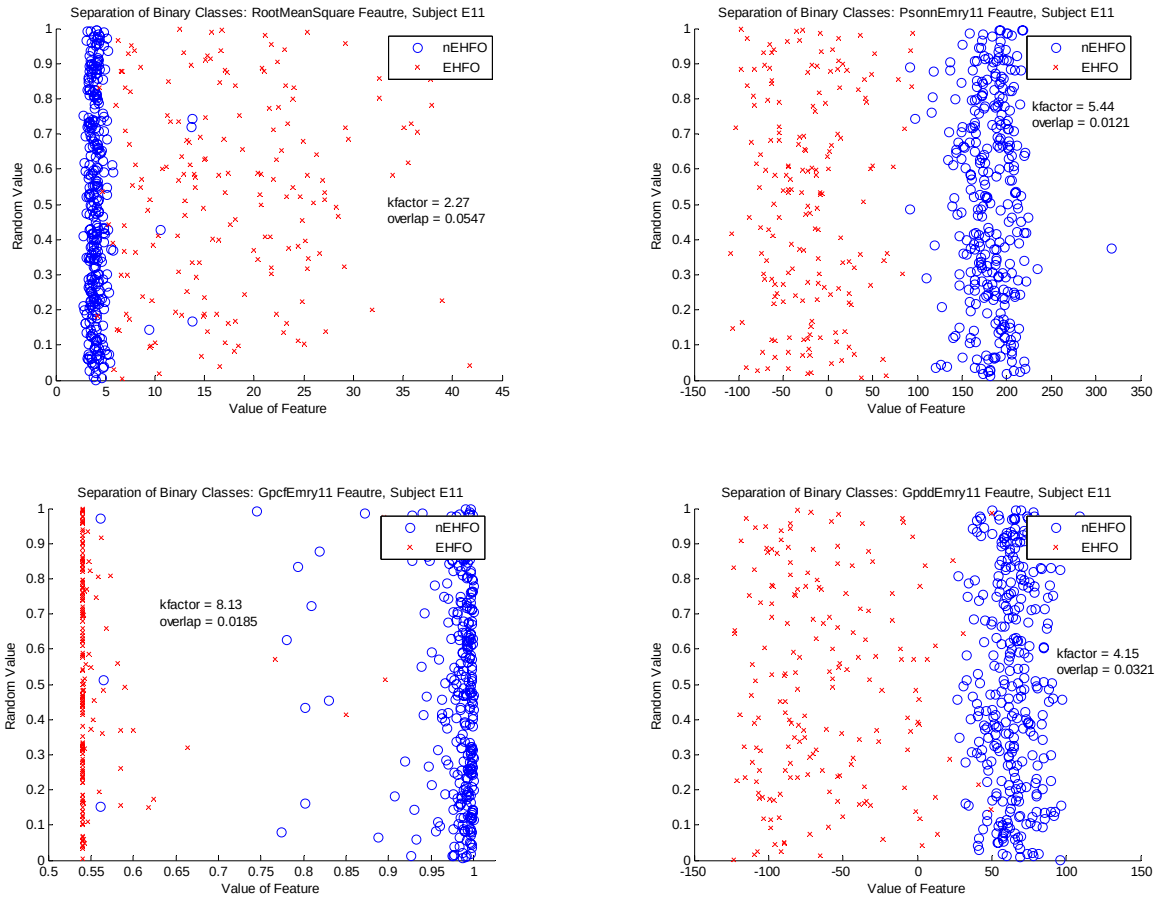


Figure 3.5-3 This figure illustrates for a subject the class-separation achieved by the presently accepted classical feature (*top left)*, genetically programmed artificial features (*bottom right, bottom left*), and a swarmed neural network artificial feature (*top right*) for detecting epileptic high frequency oscillations as represented by the feature-space. In this example, an artificial feature via GP-CF exhibited the best separation as evinced visually, corresponding with the results of comparing probability distributions but merely from another perspective.

In Figure 3.5-2, a subjective survey of the probability distributions illustrated the challenge in discerning an event as either an epileptic high frequency oscillation (EHFO) or background activity (nEHFO), while calculating the *overlap* and the *k-factor* of each distribution objectively quantified the difficulty in classification. In an alternative, but less preferred perspective of class-separation by this work, the distributions of Figure 3.5-2 essentially resembled profiles of the corresponding feature-spaces represented by Figure 3.5-3. Sharing the same x-axis of feature values, both figures revealed the generally nonlinear separation due to overlapping between observations in the binary classes (i.e., EHFO, nEHFO).

Meanwhile, Figure 3.5-5 illustrates the value of using receiver operating characteristic (ROC) curves to compare features for detecting epileptic oscillations. Although a traditional ROC curve, which stems from radar applications and terminology, would plot the probability of a correct detection against the probability of false detection to observe how each metric varies together with changes in the decision rule (e.g., threshold) of a detector, a small variant was used. That was, an ROC curve for a particular feature was generated by plotting the *sensitivity* (equal to the probability of a correct detection) versus the *specificity* (equal to one minus probability of a false detection) of a detector with a specified threshold for values of the threshold sampled over the domain of the two distributions (c.f. Figure 3.5-2) for that feature. Consequently, the characteristic curve of an ideal detector in this work coincided with a curve tracing $y = 1$ and $x = 1$ on an x-y plane. Computing the area underneath the ROC curve for each feature aptly served as an objective measure of anticipated performance to complement visual review while boding the quality for the feature.

Counting the number of correct (incorrect) detections, or true (false) positives, and the number of correct (incorrect) non-detections, or true (false) negatives permitted the computation of the sensitivity and specificity of the binary detector as described by the following expressions: $sensitivity = \frac{TP}{TP + FN}$, $specificity = \frac{TN}{TN + FP}$. These metrics, which were the abscissae and ordinates of the ROC curves, also were used in Chapter 4 to validate the performance of the detector on testing sets with EHFO and baseline activity. The *confusion matrix* in Figure 3.5-4 illustrates the space of correct and incorrect decisions for the binary classifier of the detector as well as clarifies the typical acronyms used in the formulae for performance.

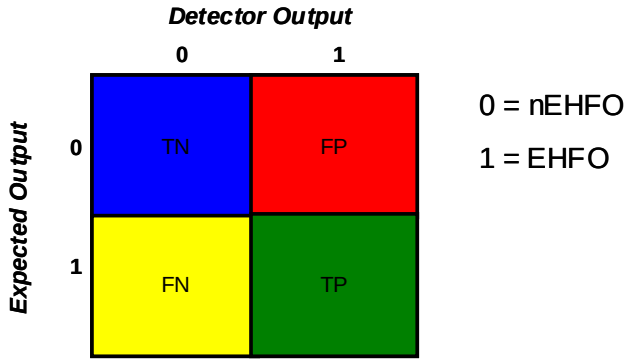


Figure 3.5-4 This figure is the confusion matrix for the detector. A true positive (TP) is a correctly detected event, a false positive (FP) is an incorrectly detected event, a false negative is a missed detection, and a true negative (TN) is a correct decision in classifying a "non-event."

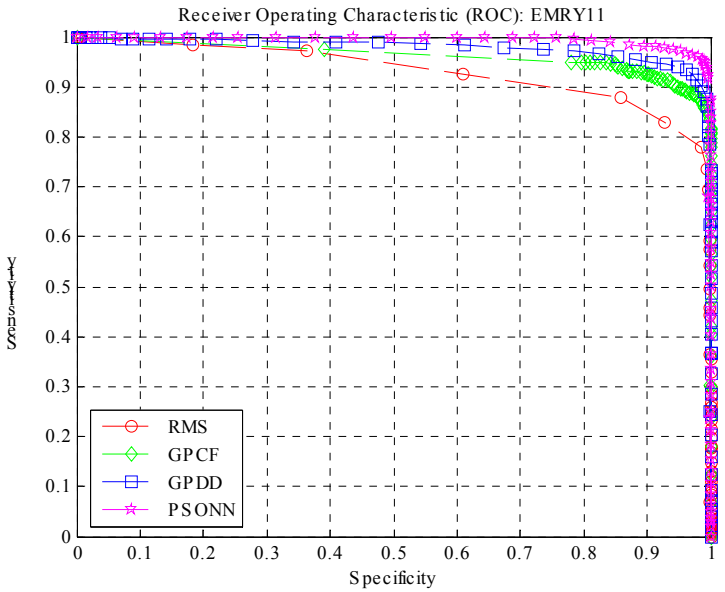


Figure 3.5-5 This figure illustrates for a subject the ROC curves achieved by the presently accepted classical feature (*red circles)*, genetically programmed artificial features (*green diamonds, blue squares*), and a swarmed neural network artificial feature (*purple pentagons*) for detecting epileptic high frequency oscillations. In this example, an artificial feature via PSO-NN exhibited the best, projected performance for a binary detector as evinced visually when an ideal ROC curve was recalled and quantitatively with the area below the ROC.

Briefly, revisiting the notion of quality features, the above figure shows that an artificial feature from a PSO-NN proved best considering the trade-off between sensitivity and selectivity of the detector for EMRY11. More specifically, the ROC for the feature from the PSO-NN was closest to an ideal characteristic curve, clearly yielding the largest area between its ROC and the *x-y* axis. This finding was consistent with the plots of the probability mass functions in Figure 3.5-2 and feature-spaces in Figure 3.5-3, considering both the k-factor and overlap for the a-feature.

Finally, a subtle point must be discussed regarding the creation of features: *over-fitting*. Traditionally, the term "over-fitting" refers to the consequence of a classifier that becomes tuned to a specific training sample rather than the underlying characteristics of the events represented by the training sample; and the classifier learns a complicated decision-boundary with "perfect" classification for the training sample but poor performance in making decisions for a novel testing sample [166]. However, by considering the optimal trade-off between performance on the training sample and the complexity of the classifier, over-fitting may be avoided with a simpler decision-rule. This notion of over-fitting in classification is illustrated in Figure 3.5-6a and Figure 3.5-6b.

Similarly, an evolutionary algorithm that was applied to creating artificial features must avoid over-fitting, or producing an a-feature that was appropriate for the training sample, but not future data as illustrated by contrasting Figure 3.5-7a and Figure 3.5-7b. For instance, the feature in Figure 3.5-7a, exhibited a high k-factor and zero overlap for the training sample, yet risked misclassification of a new testing point with a value near 2.83 that belongs to class A. More specifically, the feature in Figure 3.5-7a represented class A with a very narrow range of values, [2.85 3.00], near the outliers, [1.5 2.85], of class B and essential reflects no likelihood of the testing point being a member of class A, despite a minute difference in value of .02. This issue paralleled the tuning of a highly selective band-pass filter about its center frequency. For example, a filter the very narrow bandwidth of .1 Hz about a "perfectly" centered frequency of 426 Hz would eliminate a signal at 426.1 Hz that is practically the same as a signal at 426 Hz but slightly corrupted with frequency-modulating noise. Broadening the bandwidth of the

filter to perhaps 2 Hz—depending on the application—reasonably accommodates the error in the frequency due to the noise. Likewise, broadening the range of values over which the feature in Figure 3.5-7a characterizes class A can accommodate a reasonable variation in the true values of class A, yet still highly discriminate the two classes. Figure 3.5-7b portrays a feature that was crafted with high quality while an EA avoided over-fitting.

Two possible techniques to circumvent over-fitting in applying an EA included the incorporation of a term in the fitness that penalized complex solutions (e.g. GP trees with high depth and number of nodes) and the regulation of the stopping condition for the algorithm [166]. The latter approach was taken in this work in the following manner: 1) the structure for the PSO-NN was fixed according to [175, 182], 2) the complexity of a GP-tree was limited to a depth of eleven layers, and 3) the maximum number of generations was restricted to 35.

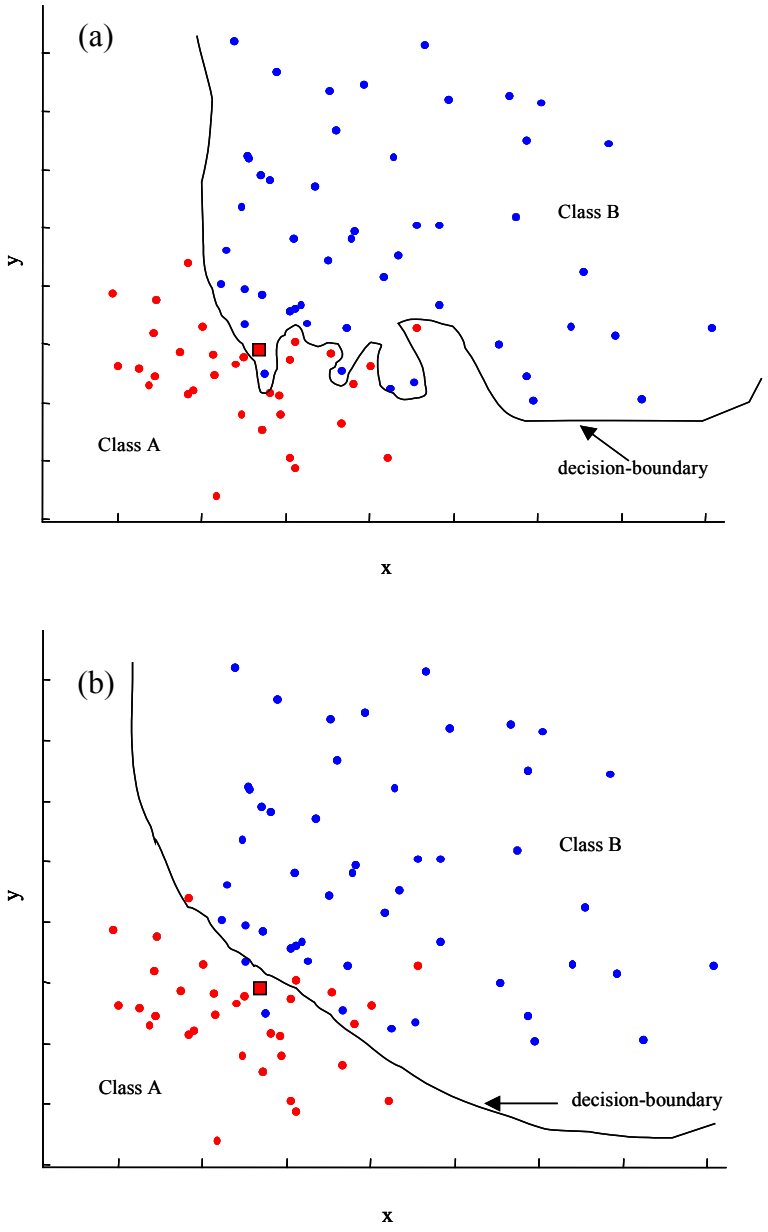


Figure 3.5-6 This figure illustrates that a complicated decision-boundary due to over-fitting (*a*) can lead to the misclassification of novel testing points (*red square*) from class A; whereas a simpler boundary (*b*) that considers the optimal trade-offs in performance of classification on a training sample may not while sacrificing better sensitivity for specificity or vice versa.

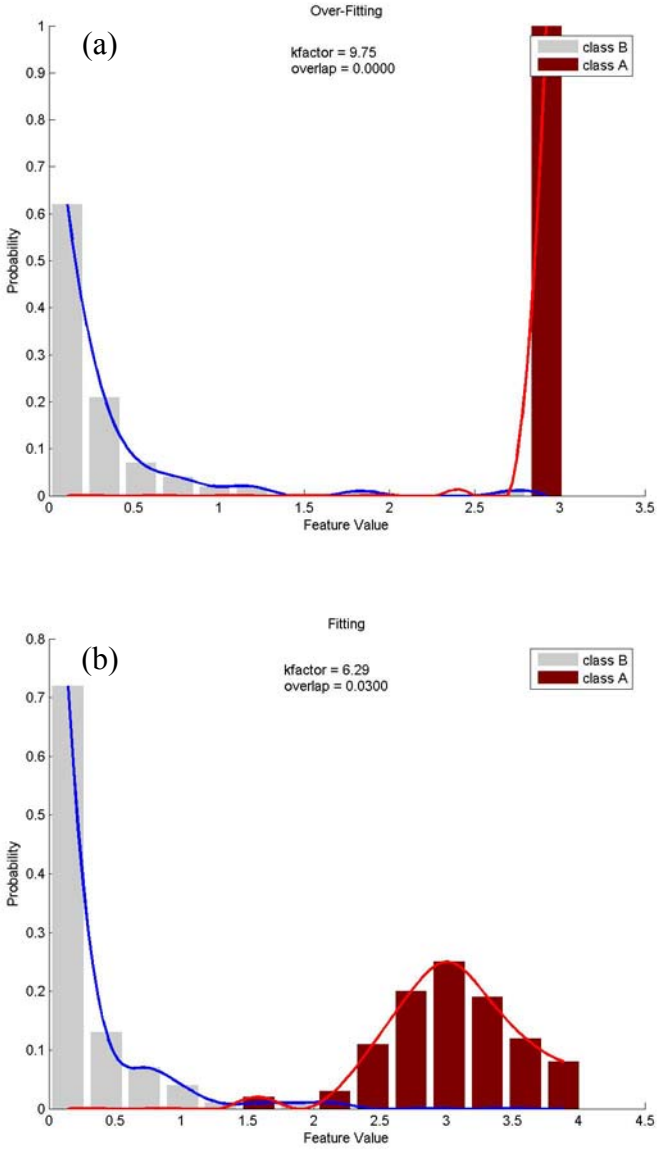


Figure 3.5-7 This figure depicts an a-feature with over-fitting (*a*) and an a-feature without over-fitting (*b*). The upper feature boasts higher separation between classes than the lower feature, but could lead to misclassification of testing points in class A that present novel feature-values (e.g., 2.8, 3.1) due to noise in the original signal from which the feature would be extracted.

3.6 Classification

Although there is no standard in detecting epileptic oscillations, popular studies have relied on assumptions about the class-separation between the prior probability distributions for the event of interest and the background activity [136, 142, 143]. For instance, detectors based upon a threshold presume that the probability density of feature values for an EHFO are always a certain distance away from the probability density of feature values for a non-EHFO [136], while more unsupervised detectors [142, 143] assumed that classes of events cluster in definite regions of the feature space. However, either parametric premise erroneously precludes the cases for homogeneous data segments, or segments with feature values from the same class pending classification. Figure 3.6-1 elucidates the flaw in a detector with a threshold or a clustering technique as a classifier.

Therefore, this work chose a classifier that made no assumptions about the underlying probability density functions or the decision-boundary between classes of the extracted feature values, presented relatively low computational burden during training, speed in classification—considering the hours of multi-channel data over which the detector would be run, and favored the least complicated approach necessary to properly define the decision-boundary [166]. Consequently, a *k*-nearest-neighbor (*k*-NN) rule was the preferred nonparametric architecture for classification. Classifiers with comparable performance using the same artificial feature, such as a probabilistic neural network (PNN) and a support vector machine (SVM), were also considered (Appendix F), but each the PNN and the SVM possessed disadvantages to a *k*-NN in computational complexity involving classification (Appendix F) and parameters for tuning the classifier

[62, 74]. Conversely, testing a *k*-NN classifier was executed through a distance metric (e.g., Euclidean, correlation) and majority rule, training simply required storing the feature values in memory, and the only parameter was the value of *k*, which controlled the error rate of the *k*-nearest-neighbor rule. In general, as *k* approached infinity, the *k*-NN rule became optimal in the sense that the error rate of the *k*-NN rule equals the Bayes error rate [166]. Figure 3.6-2 illustrates the classification of feature values using a *k*-NN.

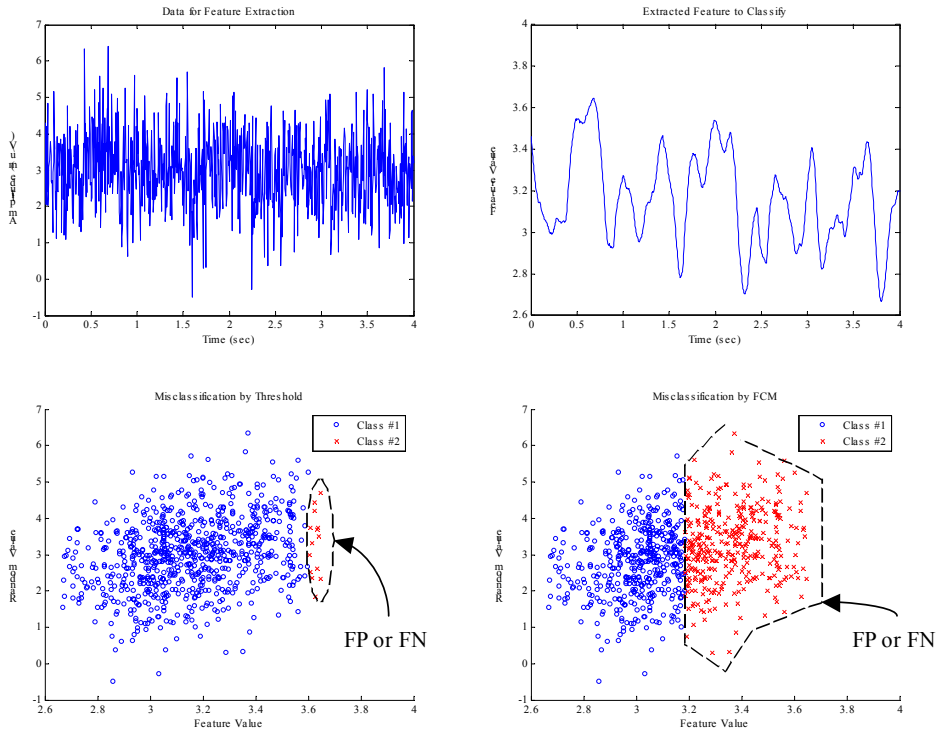


Figure 3.6-1 This figure illustrates how a detector based on a threshold (*bottom left*) or a fuzzy clustering method (*bottom right*) for a classifier will produce spurious outputs given the inputted feature values (*top right*) extracted from a homogeneous data segment (*top left*). The fuzzy clustering method was fuzzy c-means (FCM).

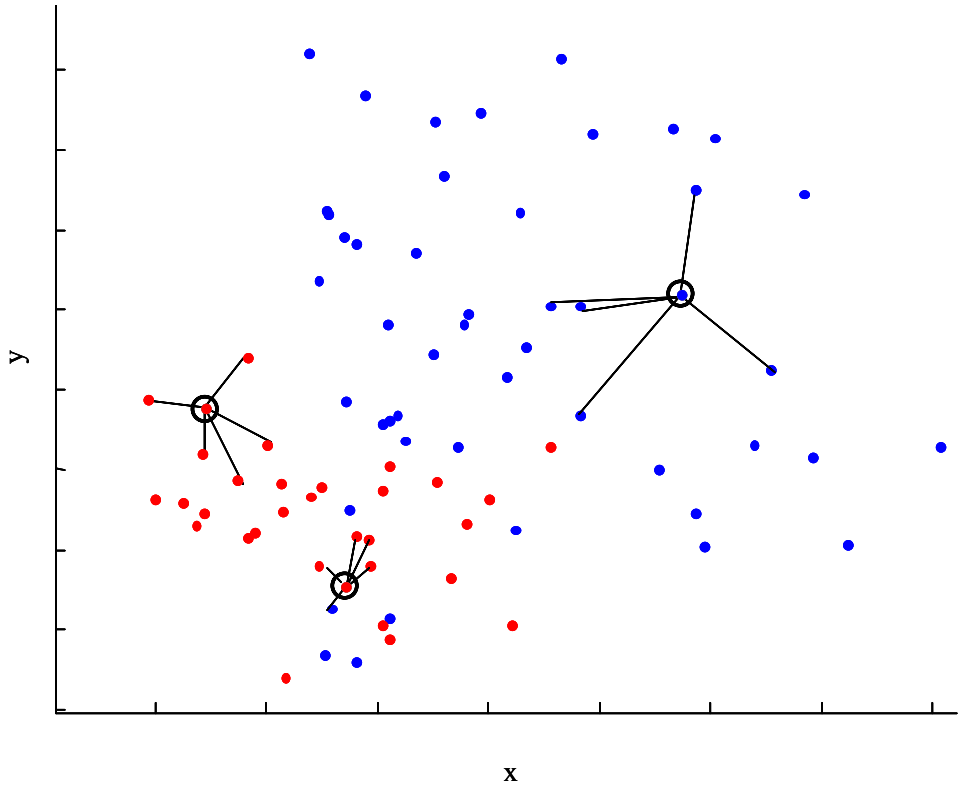


Figure 3.6-2 This figure illustrates for a two-class feature space the nonlinear, nonparametric classification of a *k*-nearest-neighbor (*k*-NN) rule with $k = 5$. The class of a test point (*black circle*) is simply determined by majority vote on the class labels corresponding to the five smallest Euclidean distances from the test point.

3.7 Detection and Validation

3.7.1 Detection

A detected event was the transition from a zero to a one (beginning time) and one to zero (ending time) at the output of the binary classifier as indicated in Figure 3.7-1. The marked times permitted measures of events for the pending FIM. The below figure illustrates the detection of high-frequency epileptiform oscillations using a malleable artificial feature.

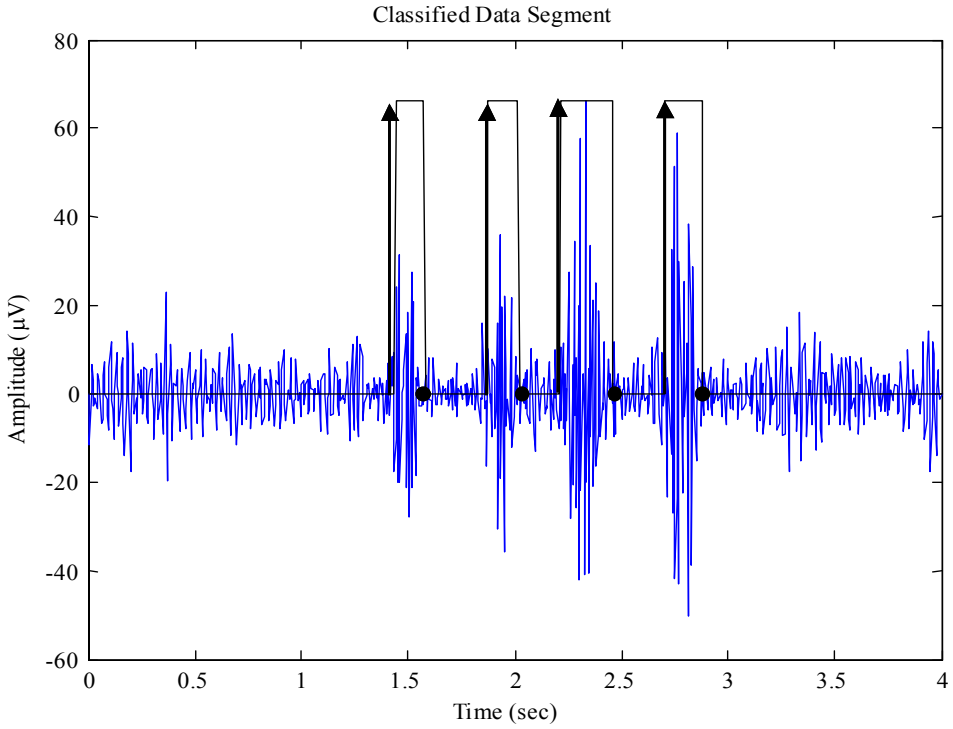


Figure 3.7-1 This figure portrays the result of the detector (*arrows and dots*) after classification (*black solid line*). More specifically, the arrows and dots represent the recorded beginning and ending times, respectively, of a detected event as recognized by the *k*-NN classifier. Essentially no detection delay is apparent because a malleable feature is used.

3.7.2 Validation

The testing data for the experiment to validate the detector included fifteen, single-channel records per subject, none of which were used to craft the artificial features or train the classifier. After the detector processed each record in the set of data, an algorithm calculated the observed sensitivity and specificity by contrasting the inventory of ground-truth markings—manually labeled by the trainer but referenced by the experienced epileptologists—against the automatic markings of the binary detector. Two aims directed the work in validating the methodology presented earlier: 1) contrast the performance of the method against a naïve detector that relied on a single physics-based feature and a threshold and served as a benchmark to investigate if the new approach to detecting epileptic oscillation improved conventional detection; 2) assess the performance of the new method relative to statistical chance.

The former aim was achieved using a two-factor analysis of variance (ANOVA) with a random effects model followed by a Tukey's multiple comparisons, when necessary, to examine evidence for a statistically significant difference between methodologies and across subjects and suggest that the performance for the new approach was better. As benchmarks to juxtapose the presented detector, this work implemented two versions of the current most successful methodology for detecting epileptic oscillations [136]: the original scheme and a variation of the original scheme that merely substitutes an optimal artificial feature for the existing classical feature, *root mean square*. The latter aim was achieved by computing the probability of the *type I error*, or decision to reject the hypothesis that the performance of the presented method was

equivalent to the performance a detector by random chance, after concatenating all testing records and performing a permutation test.

As a final note, it was important to establish the basis for estimating the metrics for the performance of the methods on the testing data. Anticipating human error in the precision of the manual markings and a possible, slight delay in triggering due to feature extraction in the detector, the value of the sliding window in feature extraction—the minimum expected duration of an epileptic oscillation—set a tolerance for any discrepancies in exact timing between the ground-truth and experimental markers. For instance, for the two patients mentioned in Section 3.3, tolerances of 160 and 500 milliseconds determined how to categorize an automatic marking as a true-positive (TP) or false-positive (FP) and an event in a testing record as a true-negative (TN) or false-negative (FN) given the manual markings. Thus, for the first (second) subject, an automatic mark beyond 160 (500) milliseconds of a manual mark was declared a FP, a manual without an automatic mark within 160 (500) milliseconds was declared a FN, a TP denoted an automatic mark within 160 (500) milliseconds of at least one manual mark, and the remaining duration of unmarked classified data consisted of TN's. Having counted the catalogued markings and events, the equations in Section 3.4 returned the sensitivity and specificity of the binary detector. Figure 3.7-2 exemplifies the abovementioned procedure for cataloguing and counting events and automatic markers in a record.

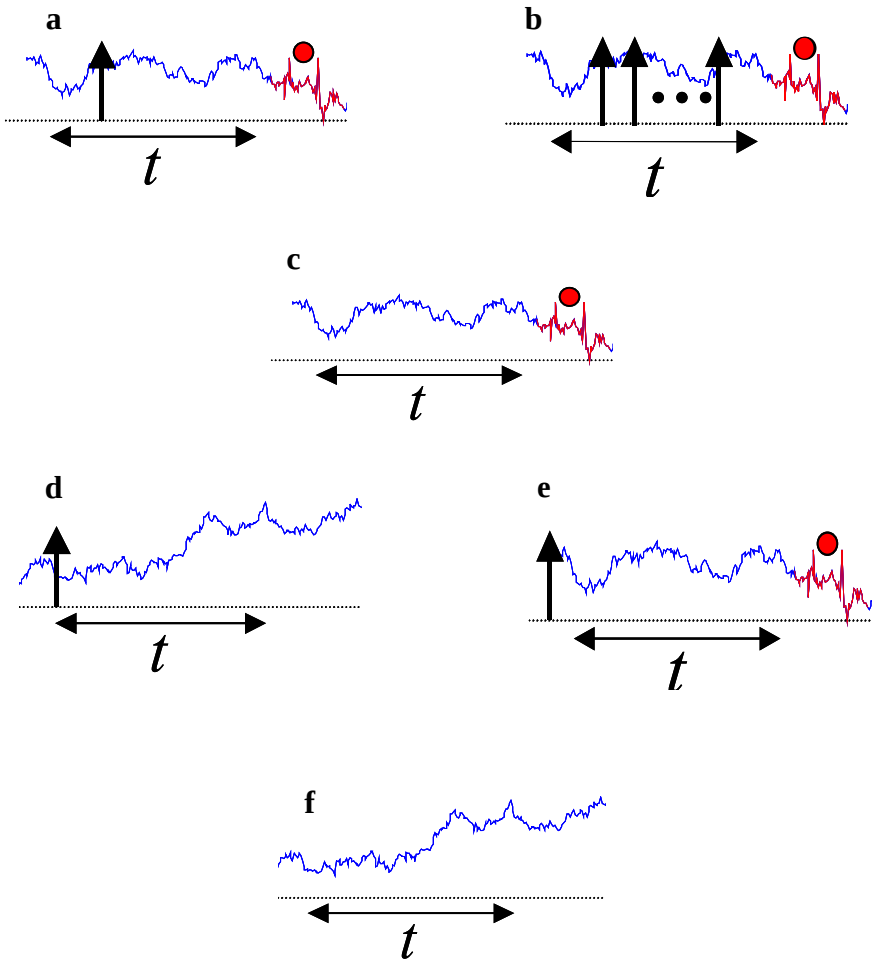


Figure 3.7-2 This figure illustrates the how automatic marks and events were categorized and tallied to compute metrics of performance for the binary detectors applied to records with known epileptic oscillations and background activity in this work. A true-positive (*a and b*) was counted when at least one mark (*arrow*) occurred within a tolerance (*t*) of a ground-truth event (*red dot*), otherwise a false-negative (*c*) was registered. A false-positive (*d and e*) denoted a mark without a ground-truth event within the tolerance, while a true-negative signified a period of tolerance without marks or events.

3.8 Mining Frequent Itemsets

To investigate if epileptic oscillations possessed any potential for pinpointing epileptogenic regions, an algorithm for FIM processed long-duration, multi-channel IEEG records prior to ictal onset after executing the detector described in the previous sections. The records were continuous IEEG with known times and locations of seizure onsets. Figure 3.8-2 illustrates the approach to reliably cluster electrodes with common concentrations of epileptic oscillations.

Using the structure and terminology for FIM from Appendix E.3, each record was segmented into one-minute intervals, where T varied between 20 and 120 minutes. All detected events within each interval per *item* contributed to a statistical sample of EHFO's that was transformed into a statistical sample of measurements according to a chosen feature (i.e., energy). For all items and all intervals, a statistic (e.g. mean, median) of each sample of measures produced a "point-measure," which represented the spatial-temporal item. Each interval of items was converted to an *n-itemset* by retaining items with a point-measure exceeding the 90^{th} percentile of all point-measures in the interval, yielding a *database* for each record. The threshold simultaneously served two purposes: 1) to account for any false alarms from the detector and 2) to only focus on rare items in connection with the feature.

Subsequently, the FIM mined each database to discover a maximal *λ-frequent* itemset at a given threshold of *support* between 0 and 1, exclusively. In this application, the support was a very subjective value, indicating the number of occurrences of a grouping of electrodes per minute. Unless otherwise stated, the support was .25, or one cluster every four minutes, for experimental results.

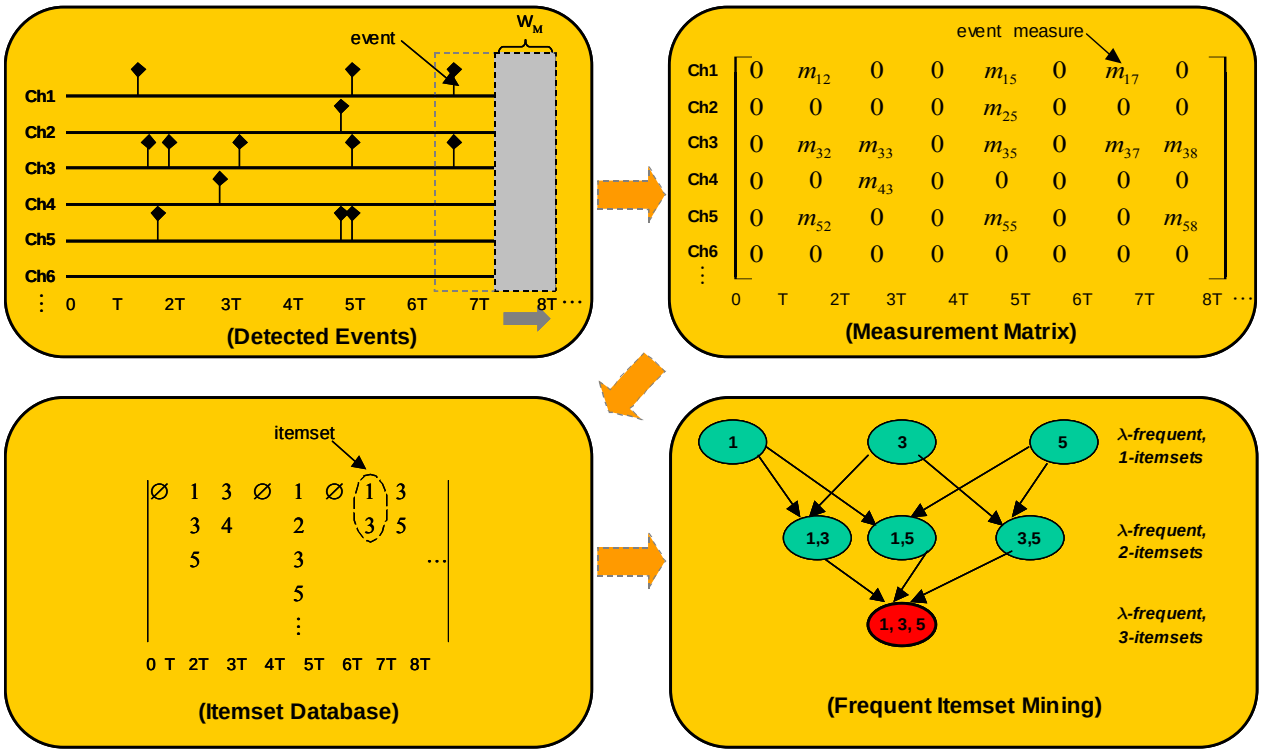


Figure 3.8-1 This figure illustrates the application of frequent itemset mining (FIM) to continuous, long-duration multi-channel IEEG with epileptic oscillations preceding seizure onsets but distributed across electrodes. The results of a binary detector (upper left) are transformed to a matrix of measures given a specific feature. This matrix becomes the input to a procedure that computes subsets of interesting channels over time. Finally, APRIORI mines the time-varying subsets to output the most frequently occurring subset, which may be of clinical significance upon review. In this example, the threshold (λ) of support (minimum rate of occurrence) for a frequent itemset (reproducible subset of channels) equals 2/9, and the maximal λ-frequent itemset, or most reproducible subset of channels at the level of support, is {1,3,5}.

Chapter 4: Experimental Results

The following four major experimental questions surfaced in contributing a comprehensive methodology to detect and analyze epileptic high frequency oscillations, which were hypothesized as critical events in mapping pathological networks to understand and counteract mechanisms in an epileptic brain:

1. How can the quality of a feature (classical or artificial) be reliably assessed, and how do artificial features compare against the standard classical feature for detecting EHFO's in terms of quality?
2. What is the performance of a detector framed in Chapter 4 with a quality feature?
3. Having a reliable detector of a particular epileptic oscillation, are reproducible spatial clusters evident in interictal records with long-duration, multi-channel recordings of EEG?
4. What are the characteristics of the epileptic oscillations, and can a model duplicate and explain those attributes?

This chapter provides details on the results of experiments that were framed to answer the above questions.

4.1 Evaluation of Feature Quality

Section 3.5 presents a means to evaluate the quality of a feature using ROC curves to complement probability mass functions or scatter-plots of the feature-space for a single subject. This section applies this technique to the remaining six patients of this work to demonstrate that an artificial feature created via an evolutionary algorithm as a more suitable feature over root mean square for detecting an epileptic oscillation. In addition, this section furthers the evaluation of features by incorporating measures of class-separation (e.g. k-factor) and non-linear regression analysis to estimate—with a

certain level of confidence—the quality of a future a-feature. That is, regression is used to project the trade-off between the sensitivity and specificity of a binary detector using an a-feature that is created with a certain k-factor. For all ensuing procedures, only the training data described in Section 3.1 is used.

For seven subjects, three artificial features that were crafted using evolutionary algorithms (i.e., GP-CF, GP-DD, PSO-NN) and the RMS were selected for a hypothetical binary classifier that declared a decision given a certain threshold and the pair (one per class) of *a priori* probability mass functions per feature. For each subject-specific feature, a threshold slid across the mass functions of the training set to produce the ROC curves in Figure 4.1-1 and Figure 4.1-2, permitting the computation of the area underneath the receiver-operating characteristic (AUROC) curve, also referred to as the *area underneath the curve* (AUC). To establish confidence in the values for the AUC at the level of 95%, the above steps were repeated after *jackknifing* the original training set for creating the features 10 times, yielding ten values that were bootstrapped to construct a confidence interval. Table 4.1-1 summarizes the values of AUC for each particular feature across the seven subjects.

Table 4.1-1 This table holds the values of the AUC as a percentage for each pair of feature and subject. The numbers in bold italics indicate the highest values per subject across features. Overall, artificial features boasted better quality (higher AUC) than the classical feature, RMS despite difficult cases (*yellow highlighting, last three subjects*) in discerning classes.

	CHOP07	EMRY02	EMRY03	EMRY05	EMRY07	EMRY09	EMRY11
RMS	98.32%	95.74%	99.58%	65.70%	93.38%	92.92%	83.45%
GPCF	98.54%	***97.81%***	***99.63%***	98.91%	96.21%	***98.05%***	98.26%
GPDD	98.52%	95.13%	99.61%	99.03%	95.93%	97.10%	98.06%
PSONN	***99.26%***	97.73%	***99.63%***	***99.28%***	***97.09%***	95.16%	***98.53%***

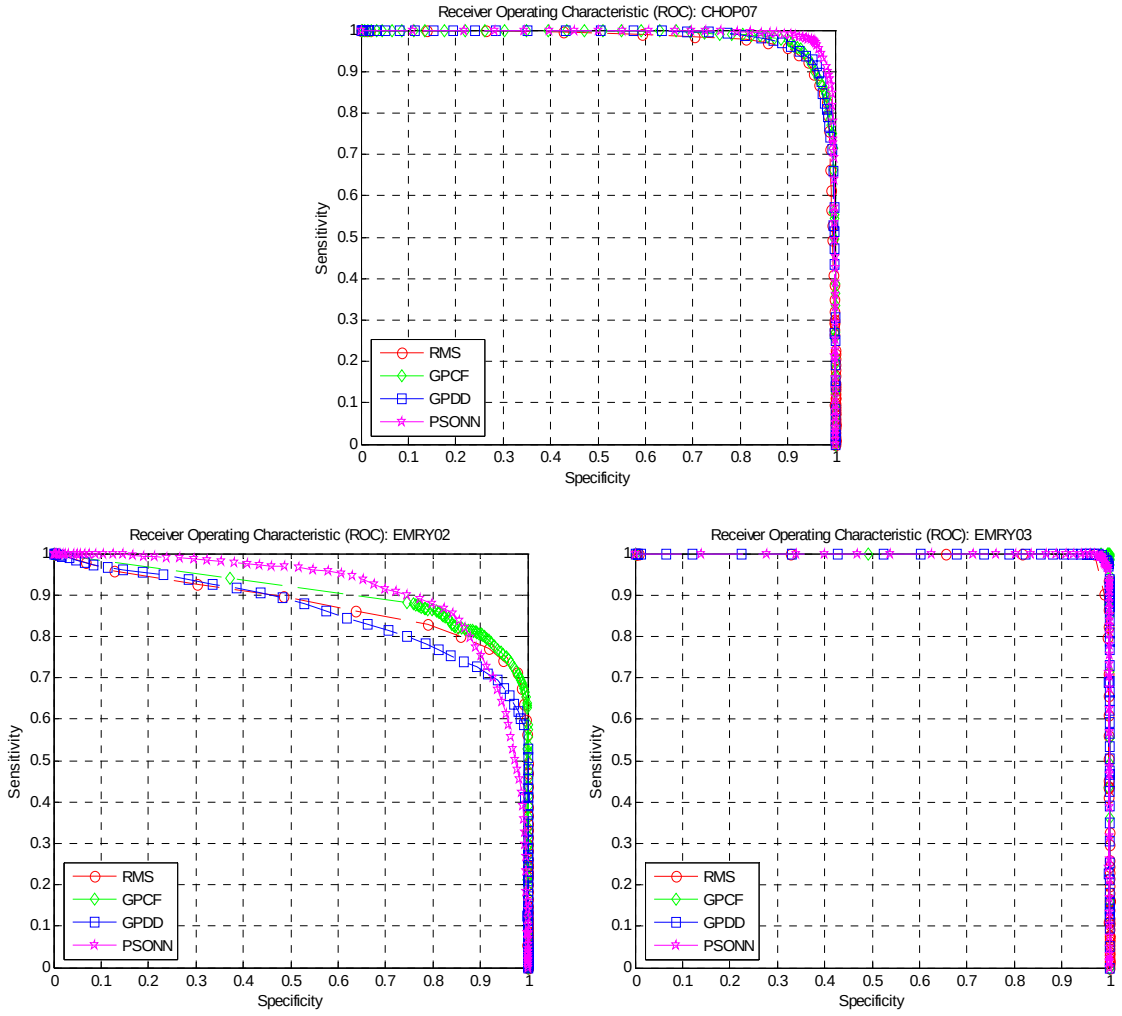


Figure 4.1-1 This figure illustrates the ROC curves (sensitivity versus specificity) for three of the seven subjects. For each subject, artificially created features by GP-CF (*green diamonds*), GP-DD (*blue squares*), PSO-NN (*purple stars*) were benchmarked against a standard feature, RMS, (*red circles*) to determine the best feature to detect epileptic oscillations. Subject EMRY03 (*bottom right*) aside, artificial features proved considerably better than RMS because characteristic curves approached an ideal ROC.

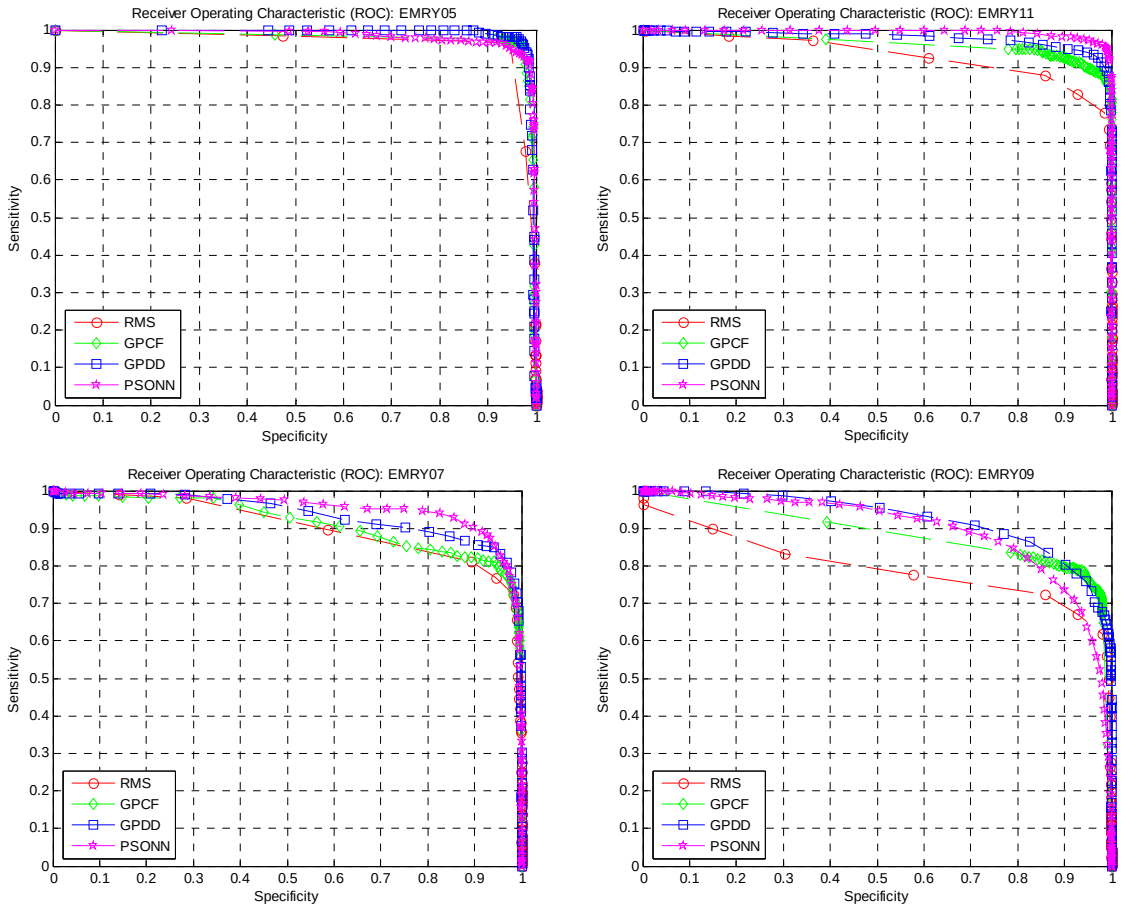


Figure 4.1-2 This figure illustrates the ROC curves (sensitivity versus specificity) for four of the seven subjects. For each subject, artificially created features by GP-CF (*green diamonds*), GP-DD (*blue squares*), PSO-NN (*purple stars*) were benchmarked against a standard feature, RMS, (*red circles*) to determine the best feature to detect epileptic oscillations. Artificial features proved considerably better than RMS because characteristic curves approached an ideal ROC.

Table 4.1-1 suggests that all artificial features outperformed RMS in terms of AUC for a binary detector of epileptic high frequency oscillations using each feature, demonstrating overall that features were crafted with better quality than simply selecting a feature based upon assumptions about physics or statistics of the events to detect. In particular, features crafted by GP-CF (i.e., EMRY02, EMRY03, EMRY09) or PSO-NN (i.e., CHOP07, EMRY03, EMRY05, EMRY07, EMRY11) relatively closely, while besting the other features. Recalling the advantages of a malleable feature noted in Section 4.4, the finding that the best feature was obtained using PSO-NN (or a high quality feature via GP-DD) was very valuable.

Figure 4.1-3 depicts another advantage of artificial features over the acknowledged classical feature to detect an EHFO: consistency in the relationship between separation of classes and anticipated performance of a binary detector and the relationship between two similar measures of class-separation. The below figure, captured an observation that artificial features across subjects consistently reflected a direct relation between the AUC of the detector and the k-factor of the density curves provided to the classifier of the detector and an inverse relation between the k-factor and the overlap—both general principles in describing performance in terms of class-separation. On the other hand, the feature RMS presented some contradiction in understanding metrics for performance and separation of classes because some subjects with low k-factor possessed high AUC and low overlap and vice versa. In sum, when creating an artificial feature, it was accurate to presume that a feature with a high k-factor would translate to high performance in classification with the *k*-NN.

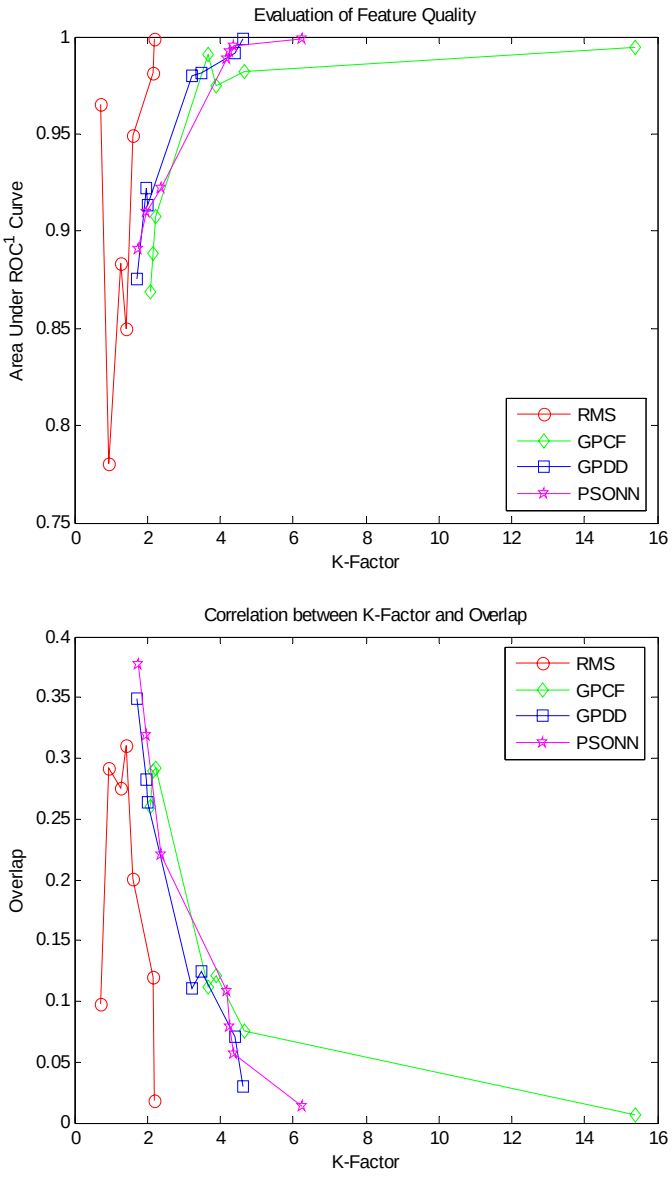


Figure 4.1-3 This figure illustrates that the expected AUC versus k-factor for the artificial features (top panel) and the overlap versus the k-factor of probability density curves for the classifier (*bottom panel*) were basically both monotonic relationships, while the currently accepted feature approximately possessed parabolic relationships for each plot. In short, this suggested a sense of inconsistency with standard feature in describing performance in terms of class-separation.

Having established a firm means to evaluate the quality of a feature and strongly supported the creation of artificial features rather than classical features for binary classification, additional work defined an apt analysis to project the quality of a different artificial feature crafted from the same training set or a new training set. In short, the following question required answering: "If a future patient-specific, artificial feature will be created, what will be the predictable quality of the feature?" Applying the theory of statistical regression to coordinates generated by computing the AUC and the k-factor of a few artificial features proved to be an adequate solution for this objective. More specifically, for each subject and evolutionary algorithm, the following steps were taken in order to answer the above question:

1. Create a feature, Φ_g, per generation, g, for several generations, G, forming a set of features $\mathrm{A} = \{\Phi_1, \Phi_2 \cdots \Phi_G\}$.
2. Choose a subset of features each with relatively dissimilar k-factors, k_i, providing a spectrum (low to high) of class-separation. Or in more formal notation, choose $\mathrm{B} \subseteq \mathrm{A} := \{\Phi_i \in \mathrm{B} \mapsto k_i \in \kappa : k_i + \varepsilon < k_{i+1} \forall i, \varepsilon > 0\}$.
3. Compute the AUC for each selected feature, v_i, producing a set of ordered pairs, $\kappa \times \upsilon = \{(k, v) : k \in \kappa \wedge v \in \upsilon\}$, with a k-factor and an AUC per pair.
4. Scatter-plot and regress the coordinates in $\kappa \times \upsilon$, computing intervals of confidence and prediction for the quality of an a-feature.

Figures 4.1-4 through 4.1-6 capture the results of this procedure, which demonstrated that evolutionary algorithms created features projecting—with 95% confidence—high performance for a binary detector.

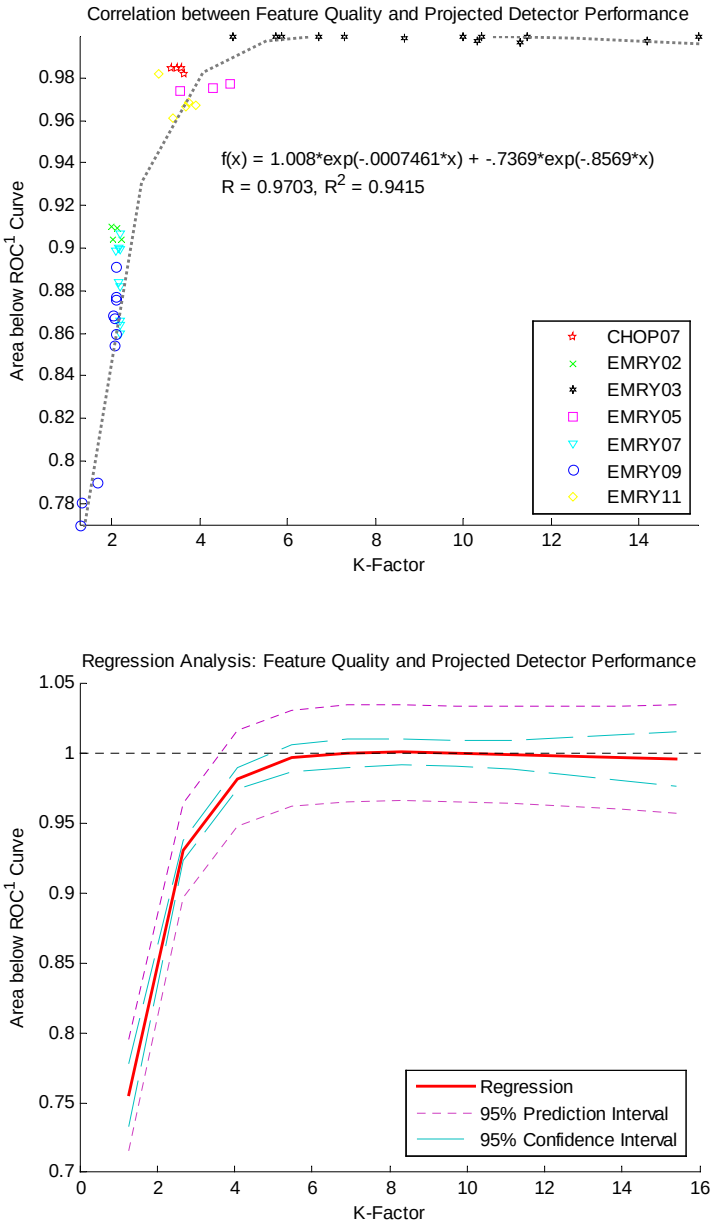


Figure 4.1-4 This figure illustrates the trend (*top panel*) and reliability (*bottom panel*) in creating a quality a-feature via GP-CF. The top panel epitomizes the fact that subjects with low (high) SNR, such as EMRY09 (EMRY03), yield a-features with low (high) k-factors and consequently low (high) quality. The bottom panel supports the idea that future, high-quality a-features can be produced with GP-CF. That is, the 95% interval for predicting new quality features by GP-CF surpasses values of AUC above .9 for k-factors greater than 3 and easily reaches the maximum of 1.0. For k-factors less than 2, marginal quality (between .7 and .85) can be achieved.

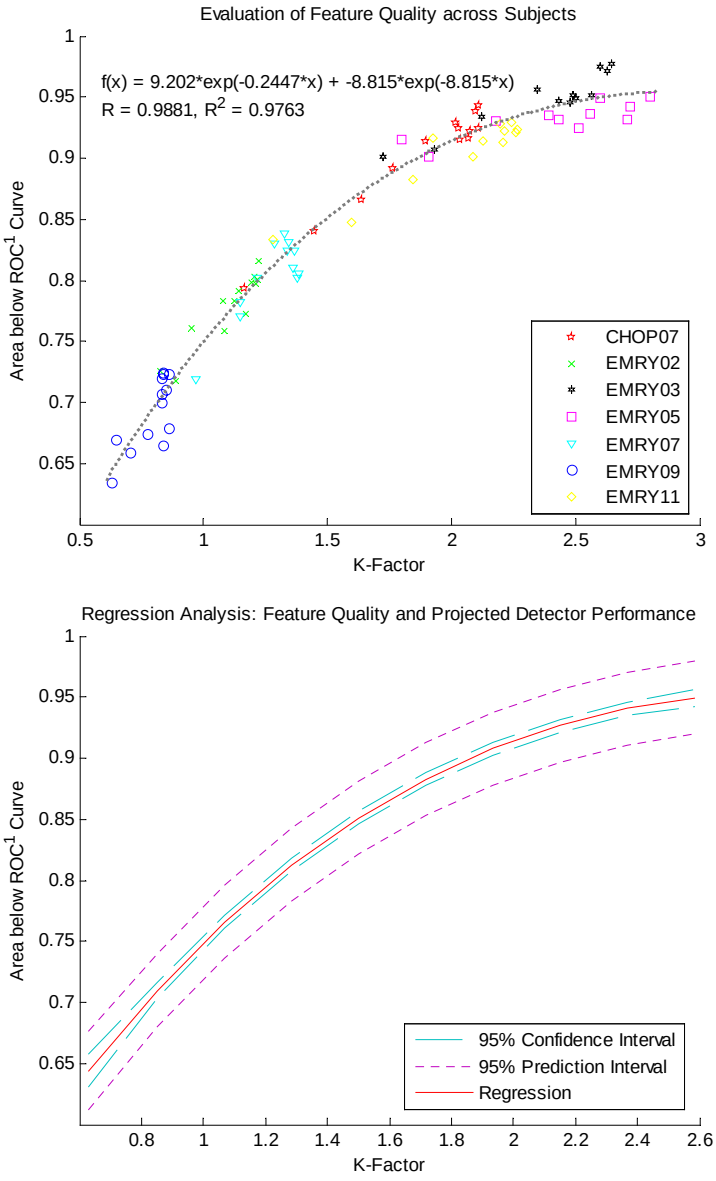


Figure 4.1-5 This figure illustrates the trend (*top panel*) and reliability (*bottom panel*) in creating a quality a-feature via GP-DD. The top panel epitomizes the fact that subjects with low (high) SNR, such as EMRY09 (EMRY03), yield a-features with low (high) k-factors and consequently low (high) quality. The bottom panel supports the idea that future, quality a-features can be produced with GP-DD. That is, the 95% interval for predicting new quality features by GP-DD surpasses values of AUC above .85 for k-factors greater than 2 but never reaches the maximum of 1.0. For k-factors less than 2, marginal quality (between .6 and .85) can be achieved. Because all k-factors are relatively low (less than 3), future creation should increase the number of generations and/or the size of the initial population before running the GP to improve quality.

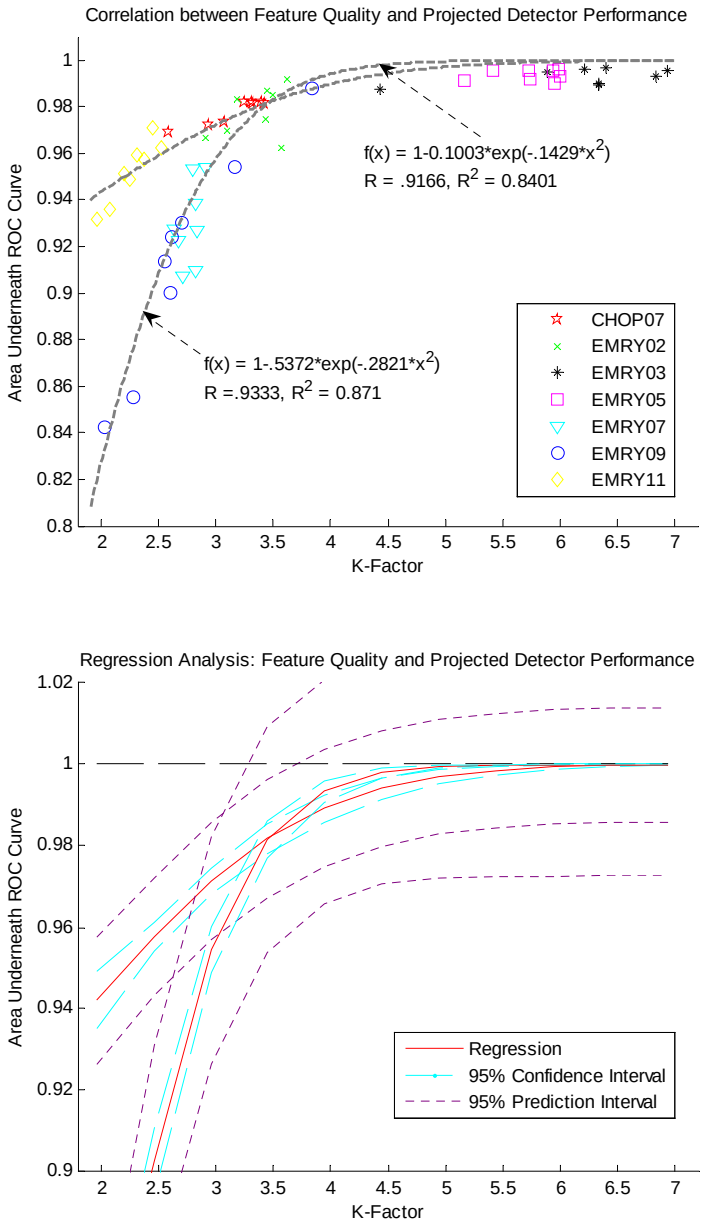


Figure 4.1-6 This figure illustrates the trend (*top panel*) and reliability (*bottom panel*) in creating a quality a-feature via PSO-NN. The top panel epitomizes the fact that subjects with low (high) SNR, such as EMRY09 (EMRY03), yield a-features with low (high) k-factors and consequently low (high) quality. The bottom panel supports the idea that future, high-quality a-features can be produced with PSO-NN. That is, the two possible 95% intervals for predicting new quality features by PSO-NN both surpass values of AUC above .9 for all k-factors and easily reaches the maximum of 1.0. The surprising finding that new high-quality features can be crafted despite cases of poor SNR further evinces the effectiveness of PSO-NN in classifying and detecting events.

4.2 Validation of the Detector

In this section, the following three methodologies were statistically tested for similarities in performance across six of the seven subjects using an ANOVA-2 and a Tukey's test for multiple comparisons:

- M1 - the "standard" methodology for detecting epileptic oscillations [136], which was based upon an unsupervised threshold of the values of a feature-series using *root mean square* (*RMS*). The threshold presumed that epileptic oscillations were statistical outliers with values in a feature-series always greater than an arbitrary number of standard deviations (e.g. 2, 5) of the mean value of the feature-series.
- M2 – a variation of M1 that addressed key concerns with the approach M1. More specifically, this method substituted an artificial feature created via PSO-NN or GP-CF for the *RMS* and set the threshold for the classifier as the threshold corresponding to the optimal point of an ROC curve for the a-feature. The idea of essentially of the latter modification was supported by Neyman-Pearson theory [166]. For curiosity, an additional modification of M1 keeping the RMS as the feature, but the threshold set using an ROC curve was also examined.
- M3 – the scheme that was presented in Chapter 4 using differentiation and band-pass filtering to enhance the time-series, an a-feature via PSO-NN or GP-CF and a *k*-nearest-neighbor rule for classification.

One subject was omitted due to difficulty in clipping a sufficient amount of baseline data for the testing set. In the ANOVA-2, the subject served as the first factor (5 degrees of freedom) and the method (3 degrees of freedom) served as the second factor, a total of 23 degrees of freedom (or the number of subjects times the number of methods minus one). The null hypothesis for each factor was that all methods to detect epileptic oscillations performed equally in terms of sensitivity and specificity. Considering each the sensitivity and the specificity as test-statistics, the ANOVA-2 returned P-values equal to zero after testing statistical similarities in performance for each factor as well as potential interactions between the two. In other words, evidence substantiated the null hypothesis as false. Since all P-values were less than 5%, a Tukey's test examined how at least one method differed from the others (the alternative hypothesis) as illustrated through the comparative box-plots in Figure 4.2-1 and Figure 4.2-2. In short, the approaches using an artificial feature improved the sensitivity of detection at the expense of lower specificity. Of the two methods using an a-feature, the one with a tuned threshold proved less taxing on specificity than the detector using a nearest-neighbor rule.

One explanation for this issue is that the less specific detectors could actually have been triggering on events missed by the human reviewers because of poor SNR and low amplitude in the EEG or events that appear physiological but are pathological. Regarding the cases of poorer than expected sensitivity in detection, such as with approach M3 with GP-CF on subject E02, it was possible that a missed event was surrounded by activity perceived by the reviewer as background though in the same band as the event—and possibly tough to discern as an event, resulting in a feature-series that rises well beyond the tolerance for the annotated time of each incorrectly detected event

and ultimately a very early trigger. If this case were true, it would necessitate better enhancement techniques than that presented, but the techniques would rely on a better understanding of the epileptic and physiological oscillations, which is unclear at the moment. In addition, from an engineering standpoint, the a-features could be improved by using an alternative metric, such as one based on the AUC perhaps, to create features more robust against the trade-off of sensitivity and specificity in a binary detector.

Table 4.1-2 captures the results of assessing the performance of a binary detector using an a-feature relative to chance after applying *permutation testing* through *Monte Carlo simulations*. Because the AUC for all binary detectors exceeded 50 percent—the value for a random detector—Table 4.1-2 merely evinced the expected outcome that no binary detector would perform less than or equal to chance. Nonetheless, for completeness in statistically validating the suggested approach to detecting epileptic oscillations, the *Monte Carlo testing* confirmed expectations with P-values less than 1% for specificity and practically zero for sensitivity.

Table 4.2-1 This table holds the values of the type I error per subject for rejecting the null hypothesis that M3 using an a-feature via PSO-NN (*a*) and GP-CF (*b*) performs equal to chance in terms of sensitivity (*SENS*) and specificity (*SPEC*). All values are below the significance level of 5%, meaning the results are unlikely to be a statistical fluke, which was expected upon recalling the ROC curves in Figure 4.1-1 and Figure 4.1-2.

	CHOP07	EMRY02	EMRY03	EMRY05	EMRY07	EMRY09
SENS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SPEC	0.00%	0.11%	0.00%	0.00%	0.01%	0.29%

(a)

	CHOP07	EMRY02	EMRY03	EMRY05	EMRY07	EMRY09
SENS	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SPEC	0.00%	0.00%	0.00%	0.04%	0.00%	0.10%

(b)

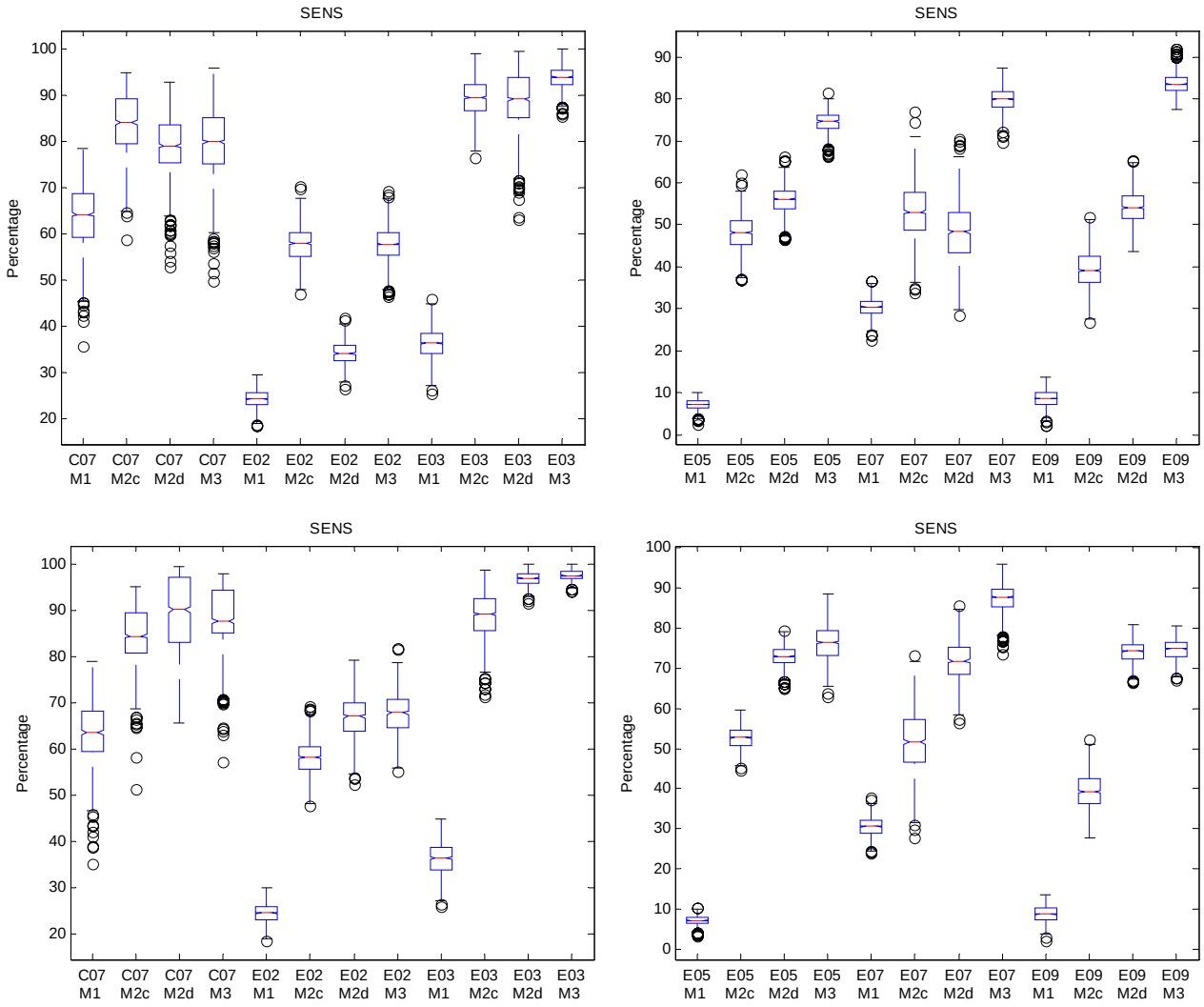


Figure 4.2-1 This figure captures the statistically significant differences in terms of sensitivity between four binary detectors (*left to right*, M1 = RMS feature with arbitrary threshold, M2c = RMS feature with tuned threshold, M2d = artificial feature with arbitrary threshold, M3 = artificial feature with a *k*-NN rule) tested on data with ground-truth markings for six subjects (*left to right*, C07, E02, E03, E05, E07, E09), where the artificial feature was created using PSO-NN (*top panels*) or GP-CF (*bottom panels*). With the exception of C07 and E02 for PSO-NN and C07 for GP-CF, M3 proved the best detector. In the latter case, M2d was the best, demonstrating that a-features were still improved detection. The notch in each box demarcates the interval of confidence in the true median metric of performance at the 95% level.

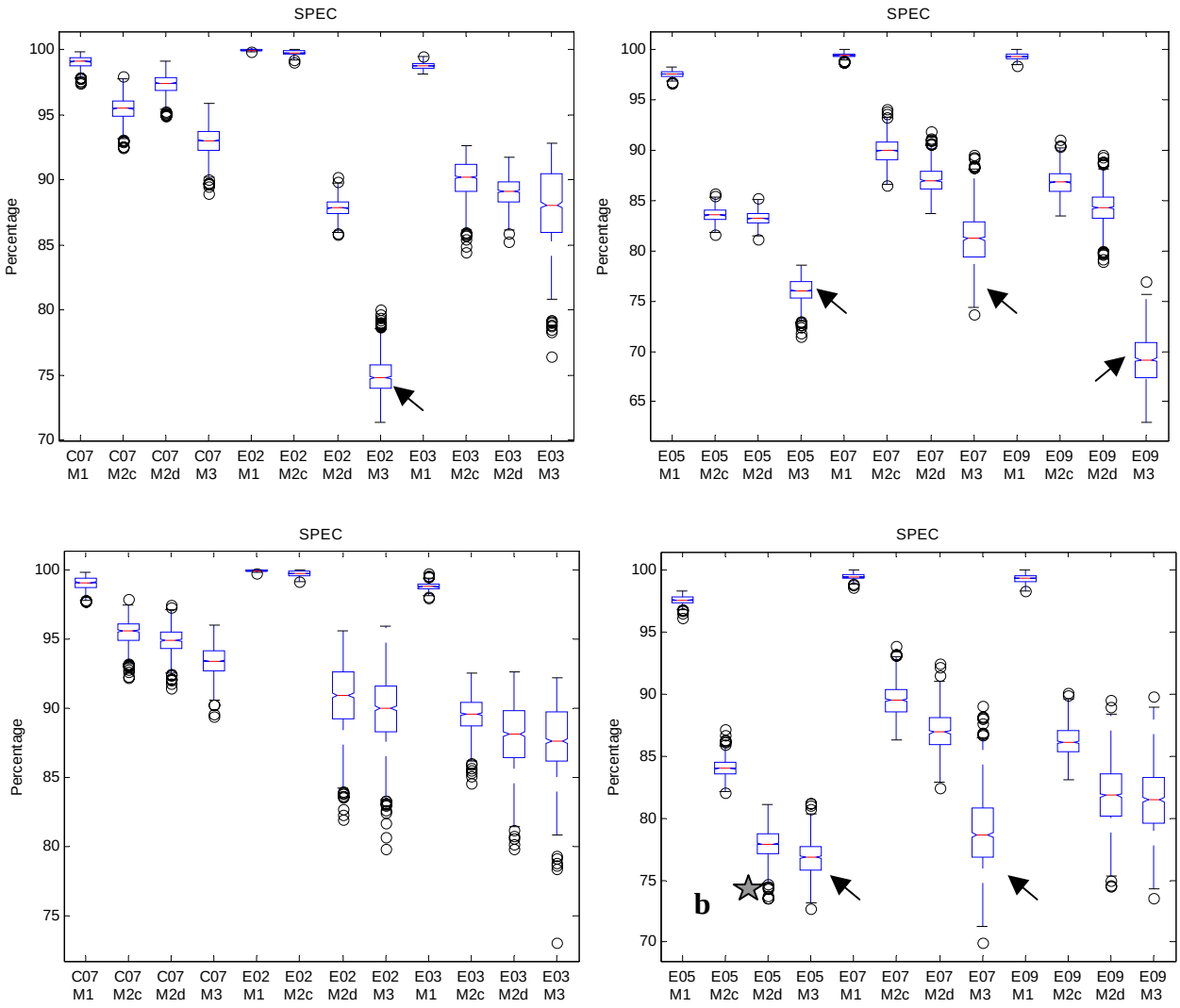


Figure 4.2-2 This figure captures the statistically significant differences in terms of specificity between four binary detectors (*left to right*, M1 = RMS feature with arbitrary threshold, M2c = RMS feature with tuned threshold, M2d = artificial feature with arbitrary threshold, M3 = artificial feature with a *k*-NN rule) tested on data with ground-truth markings for six subjects (*left to right*, C07, E02, E03, E05, E07, E09), where the artificial feature was created using PSO-NN (*top panels*) or GP-CF (*bottom panels*). Not triggering on many events—exceedingly low sensitivity—contributed to the high specificity exhibited by M1. The remaining methods demonstrated the expected trade-off between sensitivity and specificity in detection. The approach M2d, which relied on an artificial feature, achieved at least 80% specificity except in one case (*star in d*), while the other detector using an a-feature performed less well (*arrows*). The notch in each box demarcates the interval of confidence in the true median metric of performance at the 95% level.

4.3 Recognition of Putative Epileptic Networks

This section discusses the results of applying FIM to testing data of interictal records from patients with neocortical epilepsy consisting of multiple IEEG electrodes over interictal intervals ranging from several minutes to two hours after detecting a particular EHFO in the same recordings. The goal of this experiment was to investigate if the EHFO's clustered in particular regions of the brain and if any of those regions coincided with epileptic networks (e.g., the EZ, the SOZ) in each subject.

Considering the results in Section 4.2, the detector using the subject-specific a-features crafted by GP-CF and a *k*-nearest-neighbor rule as a classifier was chosen to detect the epileptic oscillations. For each subject, the *k*-NN rule relied on five neighbors, the sliding window for extracting the feature was computed according to Section 4.3, and the displacement of the sliding window was arbitrarily set to one quarter the size of the sliding window. Typically, the support for the FIM was 25%—one cluster of epileptic activity every four minutes—but differed depending upon the actual prevalence of EHFO's within each record. More specifically, if the record was sparse (dense), then the support was decreased (increased). The spatial-temporal matrix contained the mean energy of the detected signatures because Worrell and others demonstrated that the average high-frequency energy from epileptic brain was significantly greater (p-value <0.05) than in the normal control brain [53]. Only measures greater than the 90^{th} percentile of all measures were retained as a database inputted to the FIM.

The following subsections (one per patient) present the results of this analysis. Section 4.3.1 provides more detailed results than the other divisions to further illustrate the utility of the FIM. Section 4.3.8 summarizes results of all subsections.

4.3.1 EMRY03 (E03)

Eight interictal records were processed for this patient. Definite seizure onsets were visible most often in electrodes RT1-6, RIT1-4, and at times in RIF1-4. Interestingly, LIT3-4 sometimes demonstrated epileptic high frequency oscillations during ictal onset. The epileptologists marking this record indicated that the epileptic focus appeared to be RT4-6. Epileptic oscillations were detected across channels over time in each record but clustered in a few sites. Moreover, the clusters were primarily located within the SOZ, supporting hypotheses that EHFO's might be useful in pinpointing the SOZ.

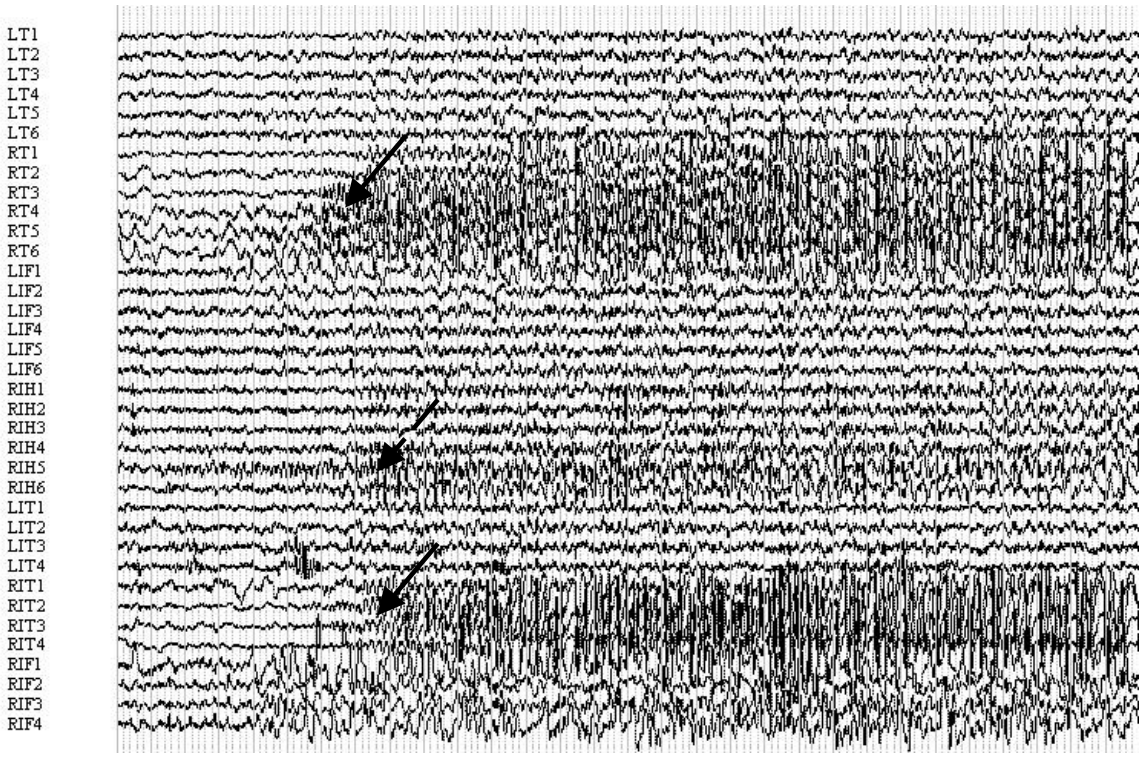


Figure 4.3-1 This figure illustrates the SOZ for E03, which typically began in RT1-6 (*arrows*) before rapidly spreading to RIT1-4 and possibly RIH5-6 (*broken arrow)*.

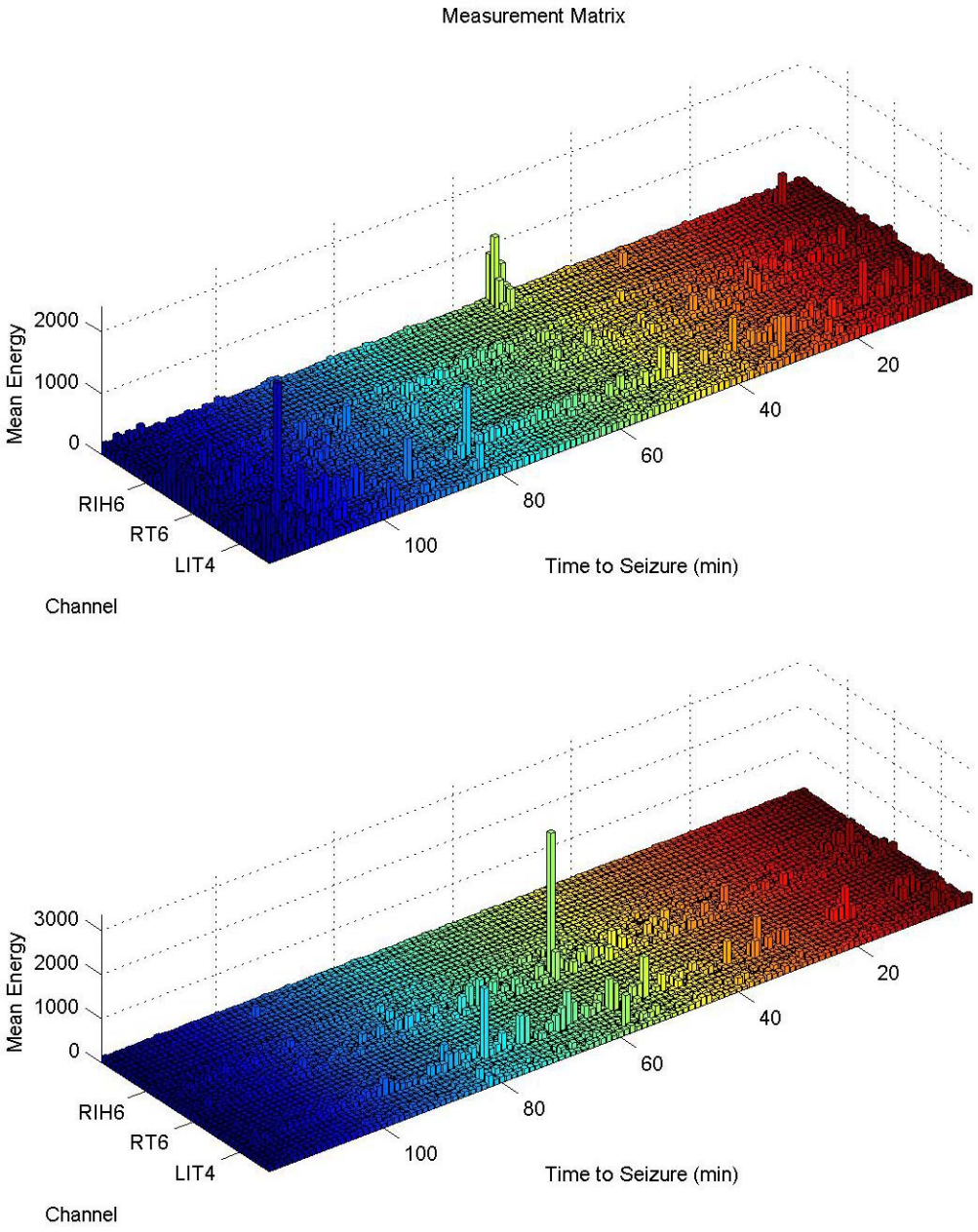


Figure 4.3-2 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject E03. Already it can be observed from the two different interictal records that certain channels (e.g., RT5, RT6, LIT4) frequently experience interictal epileptic oscillations.

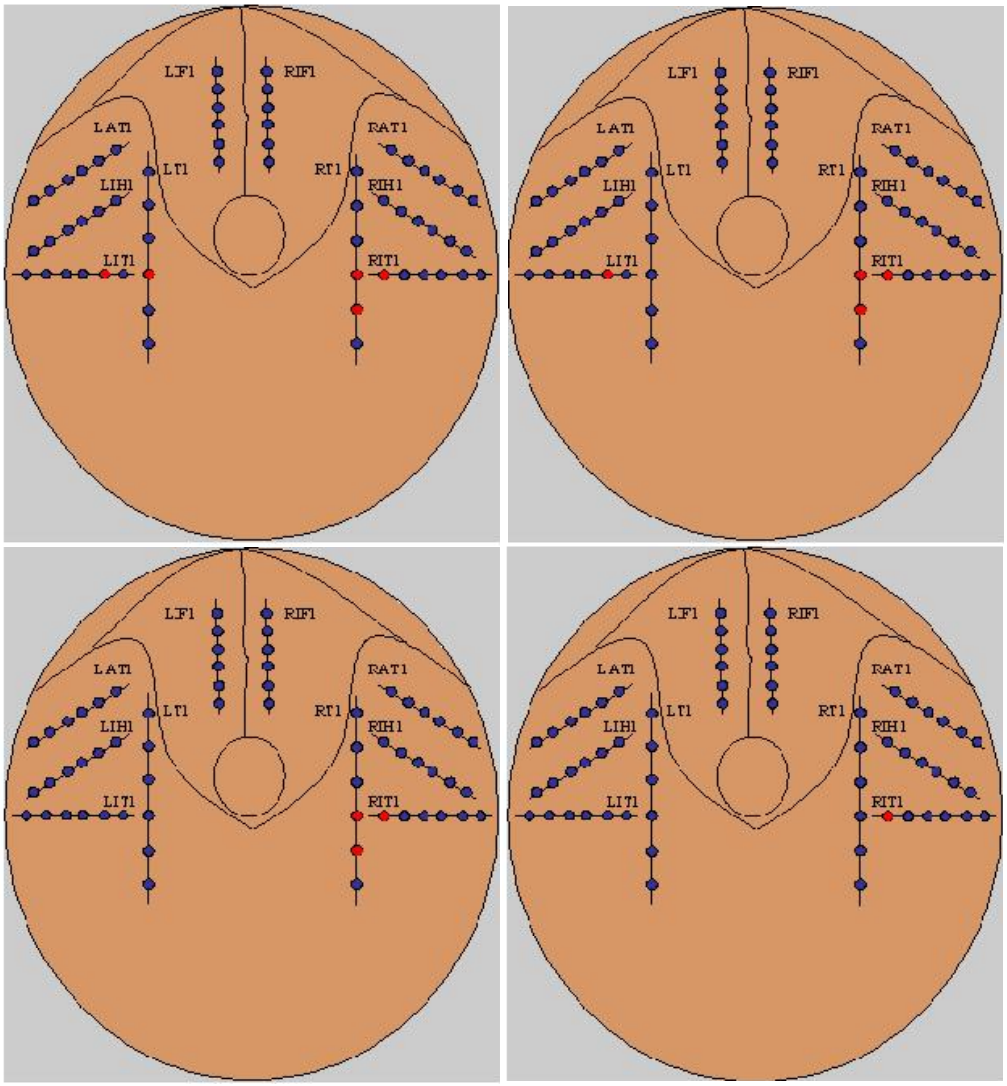


Figure 4.3-3 This figure illustrates the output to the module that statistically clusters the spatial-temporal distribution of the energy of EHFO's for E03 using FIM with four different sizes of clusters at a support of 40%. The single most frequent channel (*bottom right panel*) was RIT1, which coupled with two other channels (RT5, RT6) in the SOZ and two channels (LIT2, LT4) outside the SOZ.

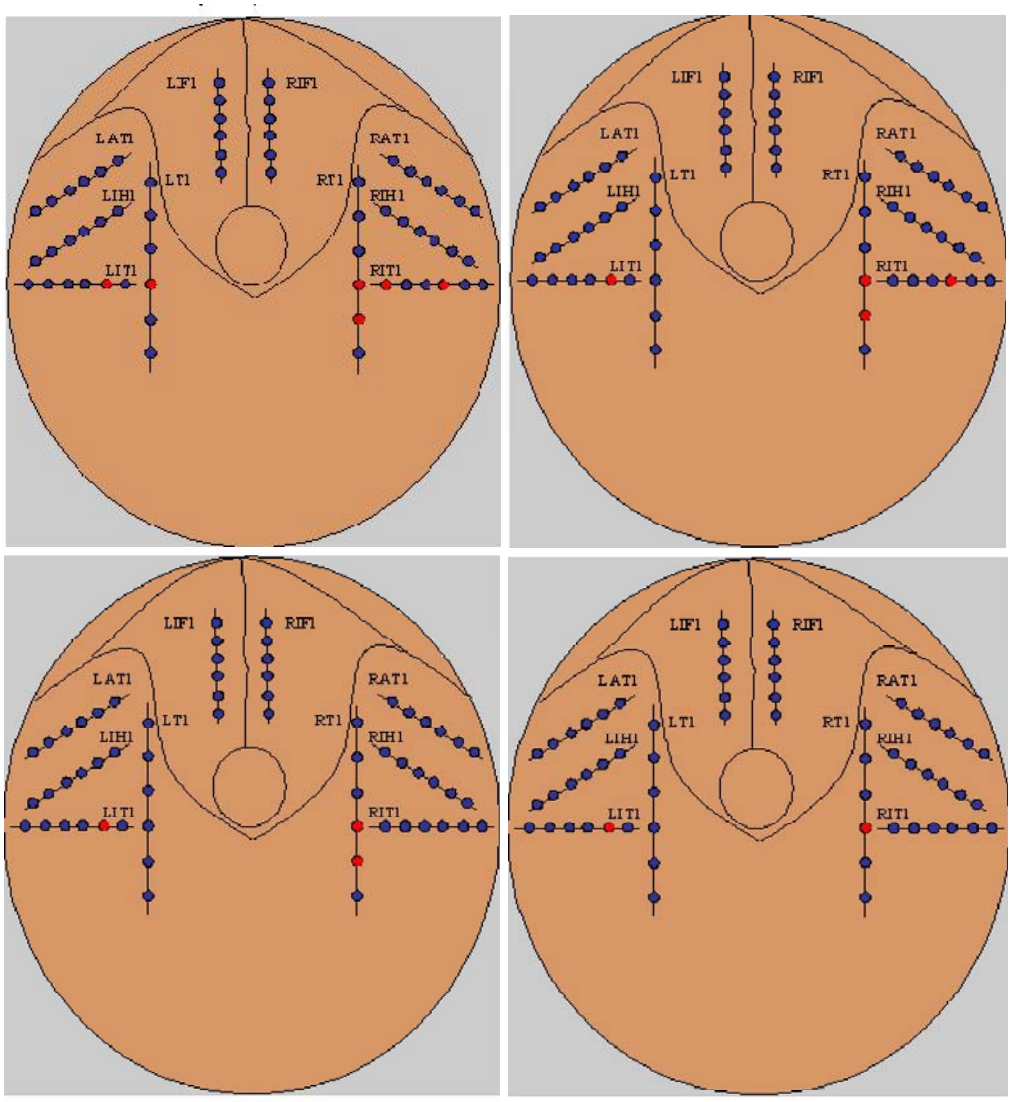


Figure 4.3-4 This figure illustrates the output to the module that statistically clusters the spatial-temporal distribution of the energy of EHFO's for E03 using FIM with four different values of support: 25% (*top left*), 30% (*top right*), 40% (*bottom left*), and 50% (*bottom right*). The support determines the coupling of frequent channels with higher levels returning smaller clusters.

4.3.2 CHOP07 (C07)

Fourteen interictal records were processed for this patient. Seizure onsets were clearly discernible in electrodes 65, 66, 67, 68, 74, 81, 82, 91, and 92 with channels such as 34, 35, and 119 sporadically exhibiting seizures but always high frequency energy mainly before the ictal onsets in the other channels. In each record, epileptic oscillations were detected only in certain channels over time and clustered in even fewer sites that were located either within or near the SOZ, supporting hypotheses that EHFO's might be useful in pinpointing the SOZ.

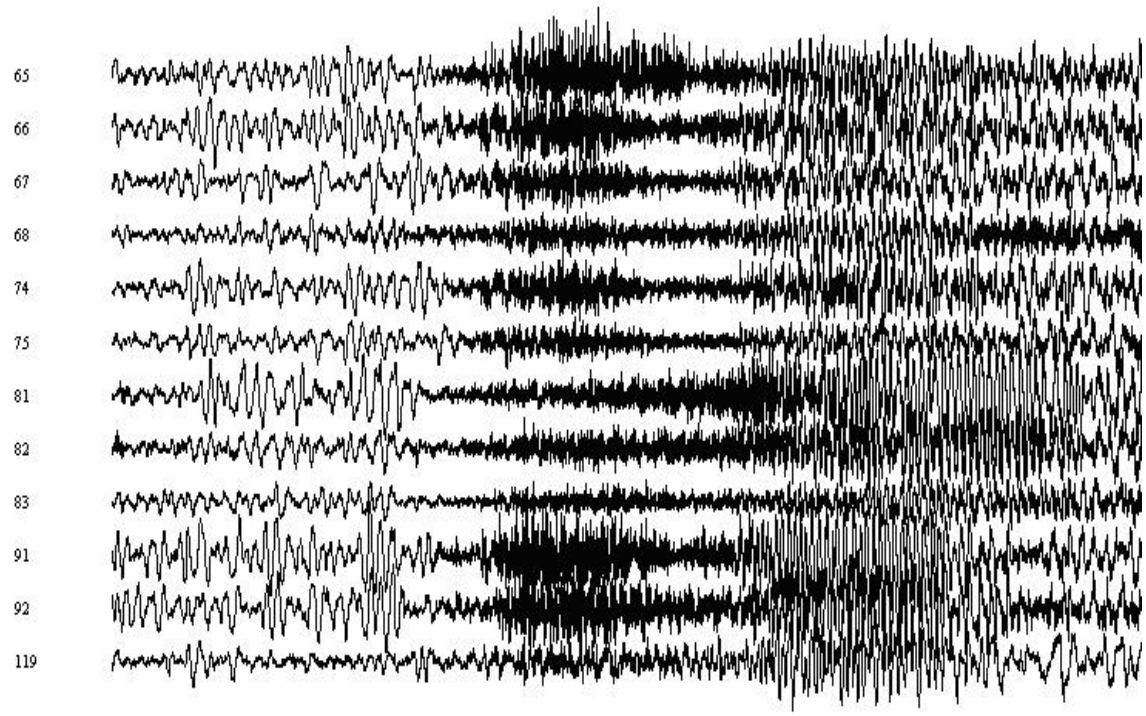


Figure 4.3-5 This figure illustrates the SOZ for C07, which typically consisted of electrodes 65, 66, 67, 68, 74, 81, 82, 91, and 92.

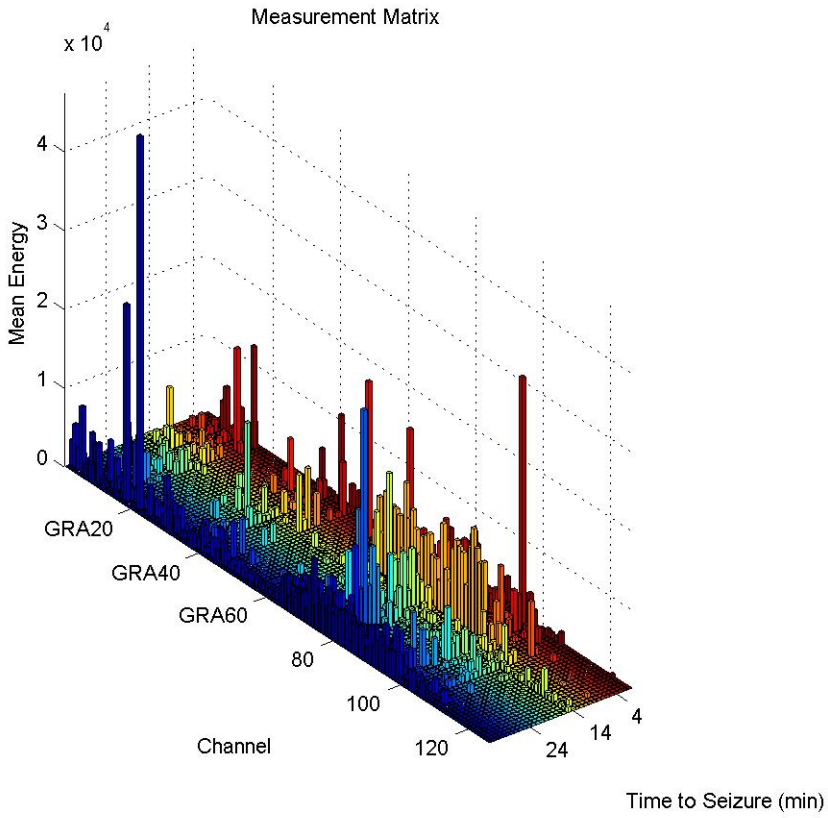


Figure 4.3-6 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject C07. It is the measure matrix for the interictal record captured in Figure 4.3-5.

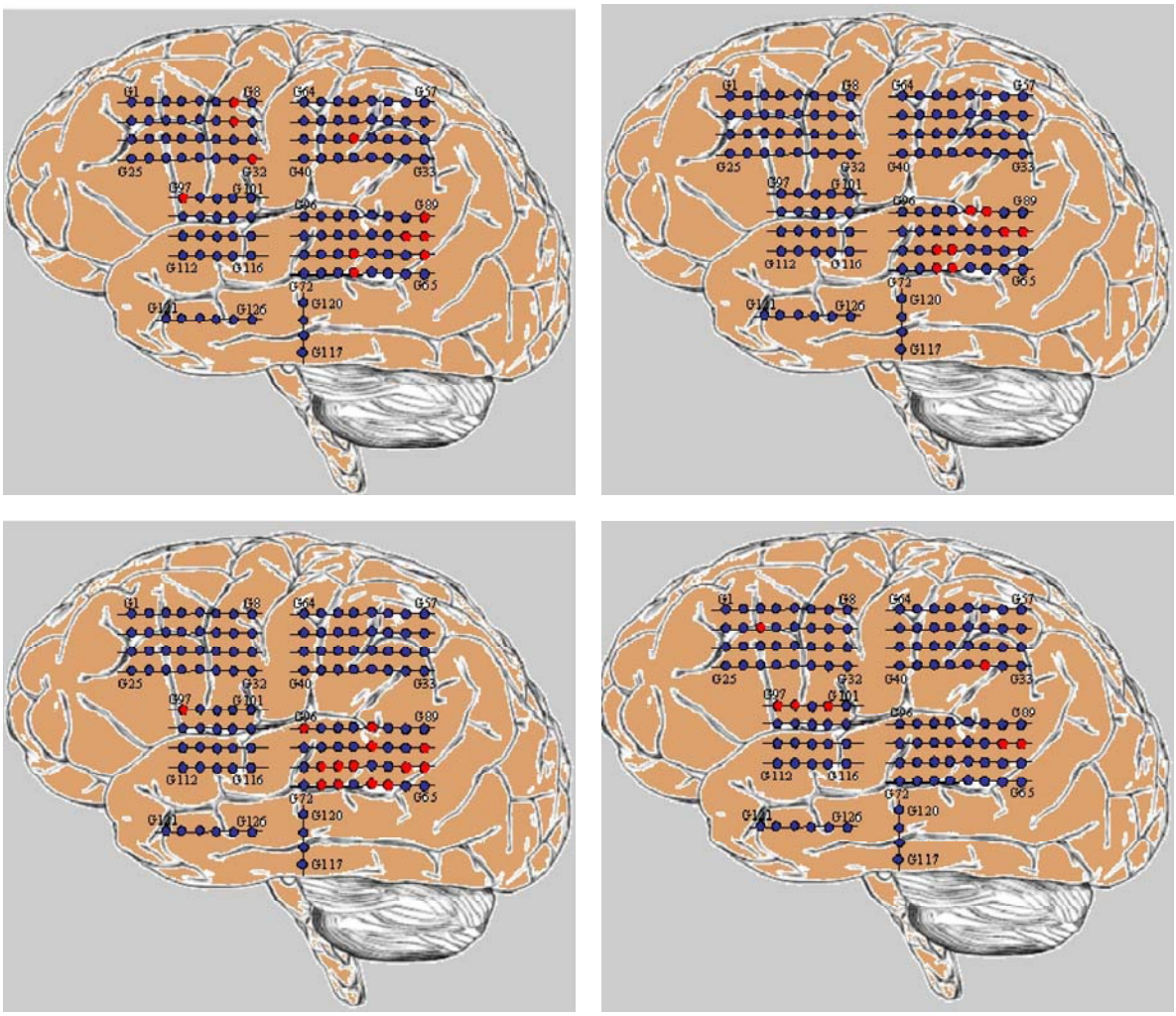


Figure 4.3-7 This figure illustrates the clustered electrodes for four different records from C07 at a support of 30%. Often the coupled channels coincided with or were near the SOZ for this subject.

4.3.3 EMRY02 (E02)

Six interictal records were processed for this patient. Ictal onsets were most often visible in electrodes LIF1-2 and RIT4-6, while at times in LIT2-4. The epileptologists who reviewed this record suggested RIT4-5 as the epileptic focus. The detection of epileptic oscillations varied each record affecting the overall results of the FIM, necessitating different thresholds across records and resulting at times in clusters outside as well as within the SOZ.

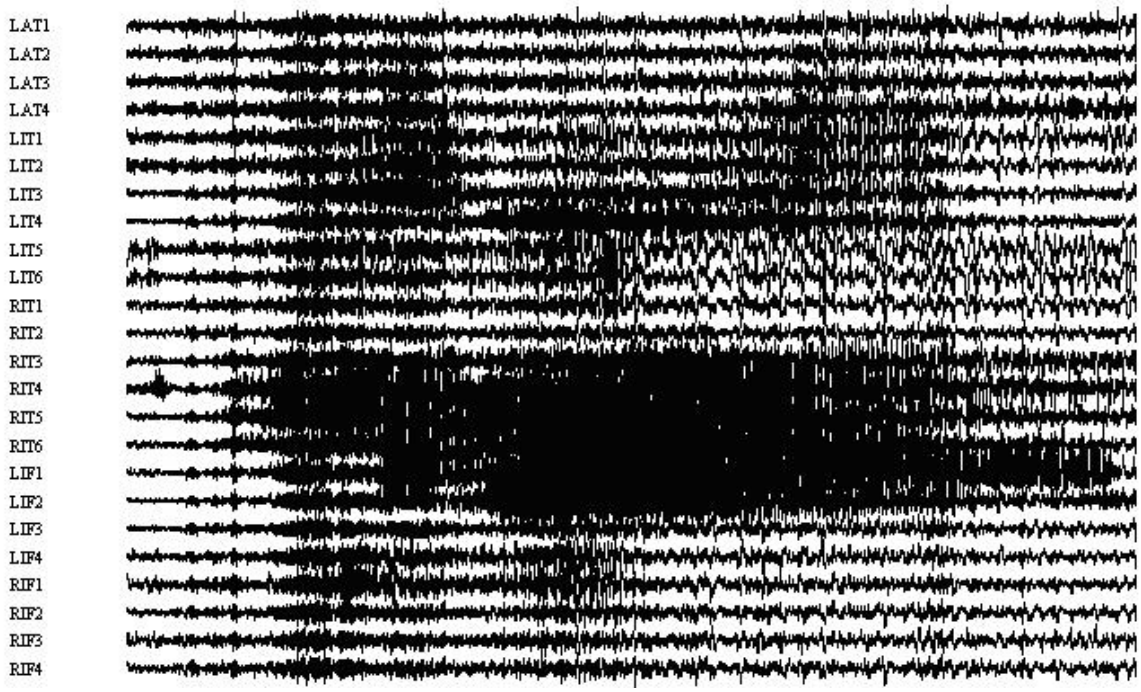


Figure 4.3-8 This figure illustrates the SOZ for E02, which typically consisted of electrodes LIF1-2, RIT4-6, and sometimes LIT2-4.

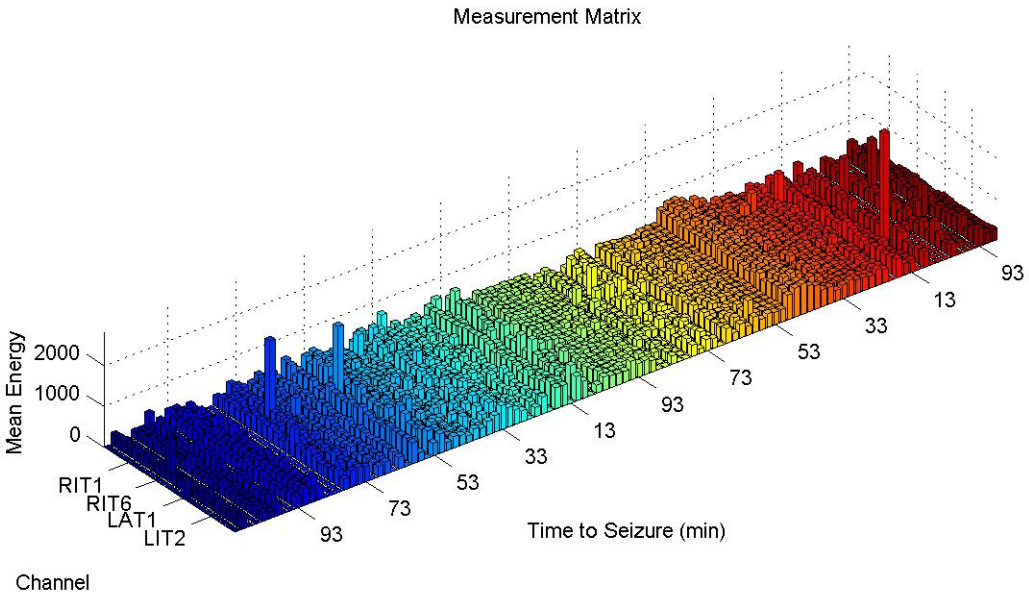


Figure 4.3-9 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject E02. It is the measure matrix for the interictal record captured in Figure 4.3-8.

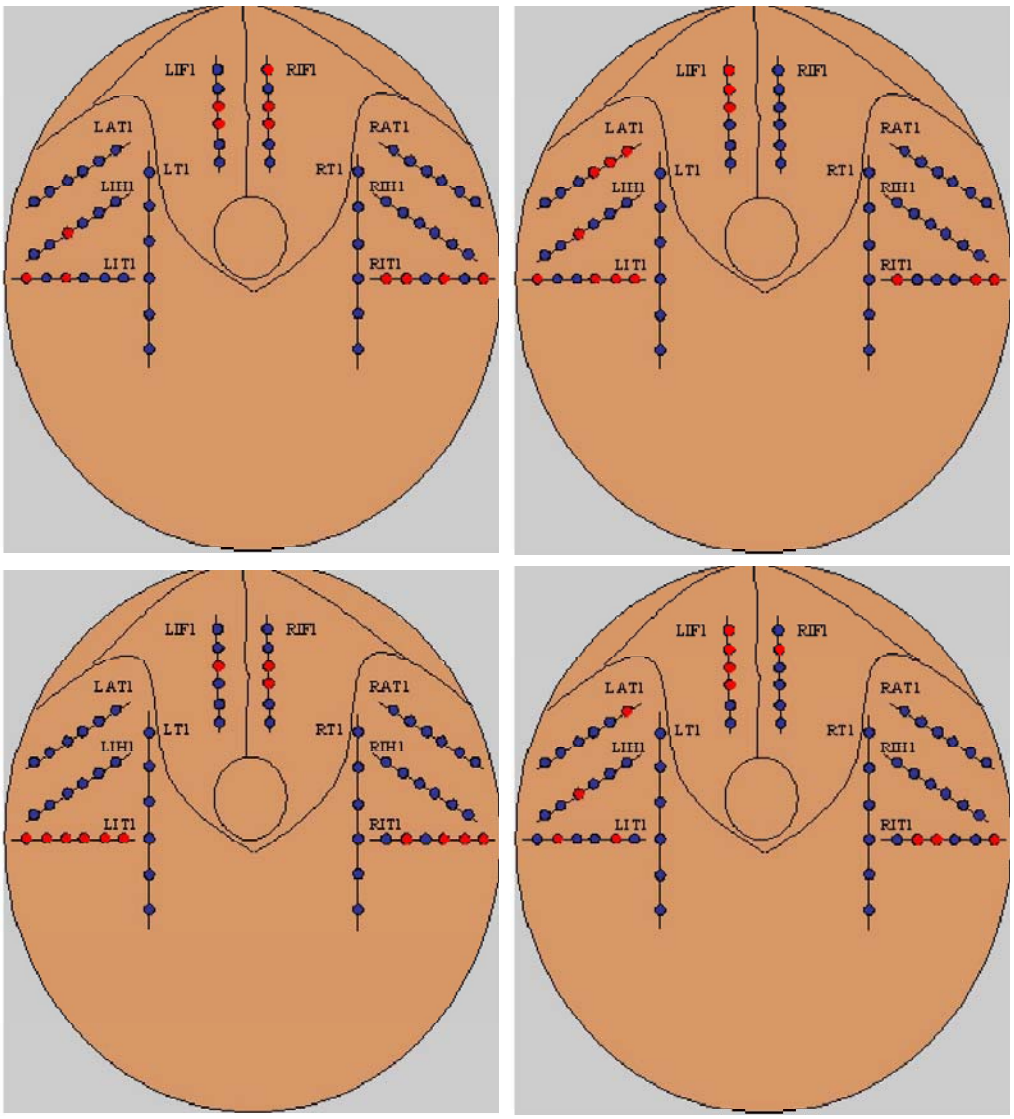


Figure 4.3-10 This figure illustrates the clustered electrodes for four different records from E02 at a four levels of support: 17% (*top left*), 65% (*top right*), 50% (*bottom left*), and 35% (*bottom right*). In some records, the coupled channels included electrodes in regions outside the SOZ (e.g., LAT, RIF).

4.3.4 EMRY05 (E05)

Four interictal records were processed for this patient. Ictal onsets were most often visible in electrodes RT2-5, RIT1-4, and RIF1-4. The epileptologists who reviewed this record selected RT2-3 and RIT1-2 as the epileptic focus. Interestingly, applying FIM to the epileptic oscillations detected across each record most consistently yielded clusters within RIF, which was still part of the SOZ, in three of the four cases.

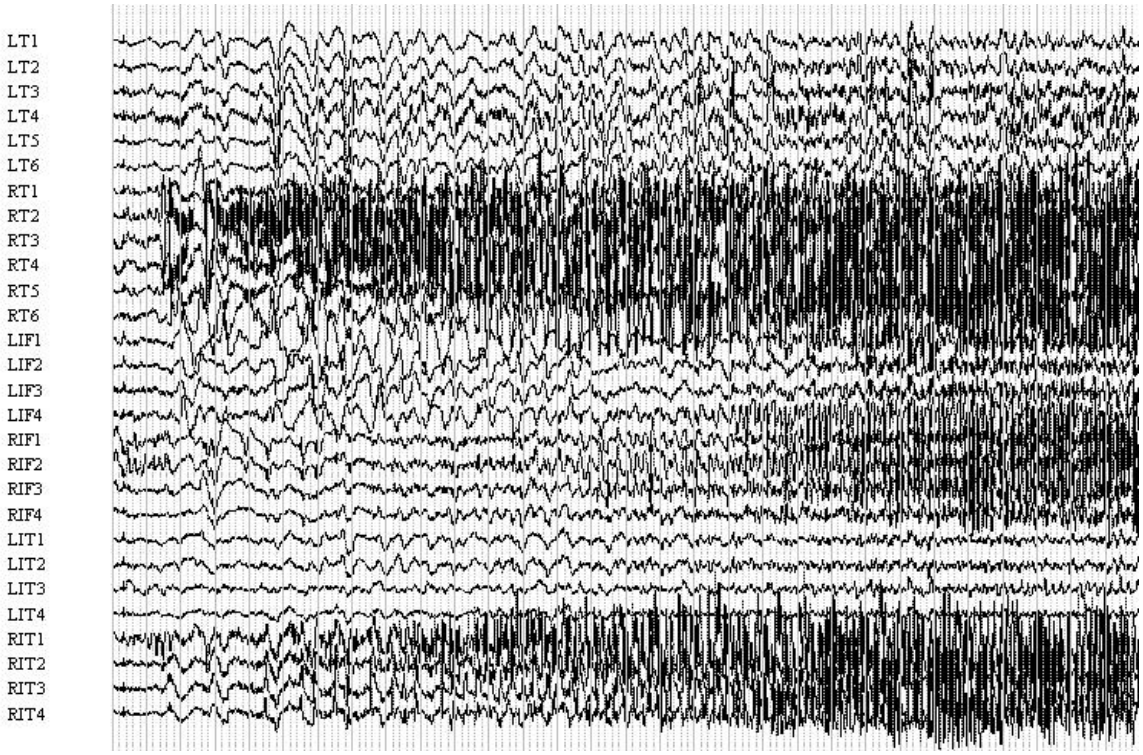


Figure 4.3-11 This figure illustrates the SOZ for E05, which typically consisted of electrodes RT2-5, RIT1-4, and sometimes RIF1-4.

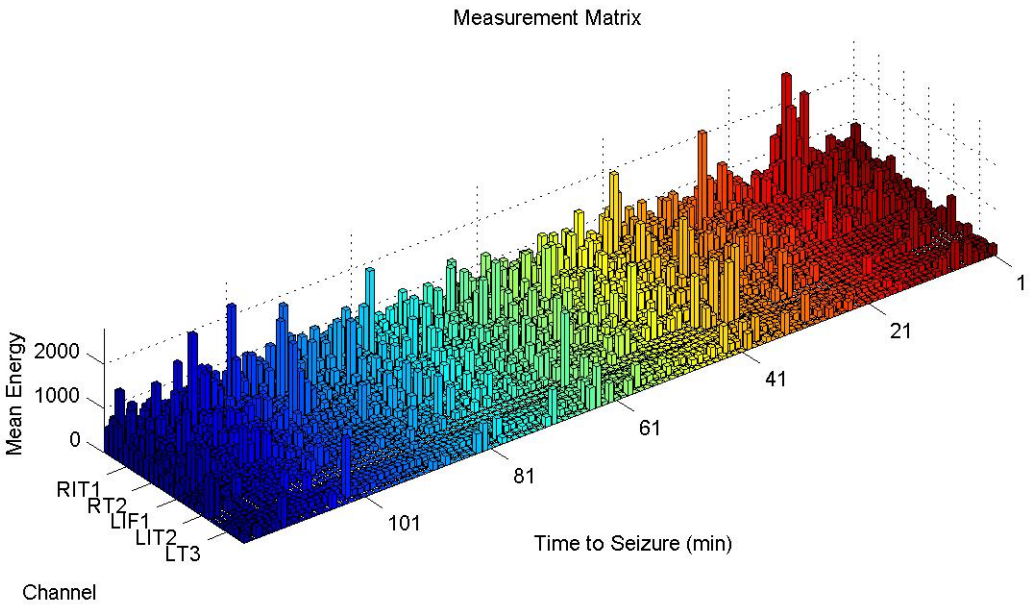


Figure 4.3-12 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject E05. It is the measure matrix for the interictal record captured in Figure 4.3-11.

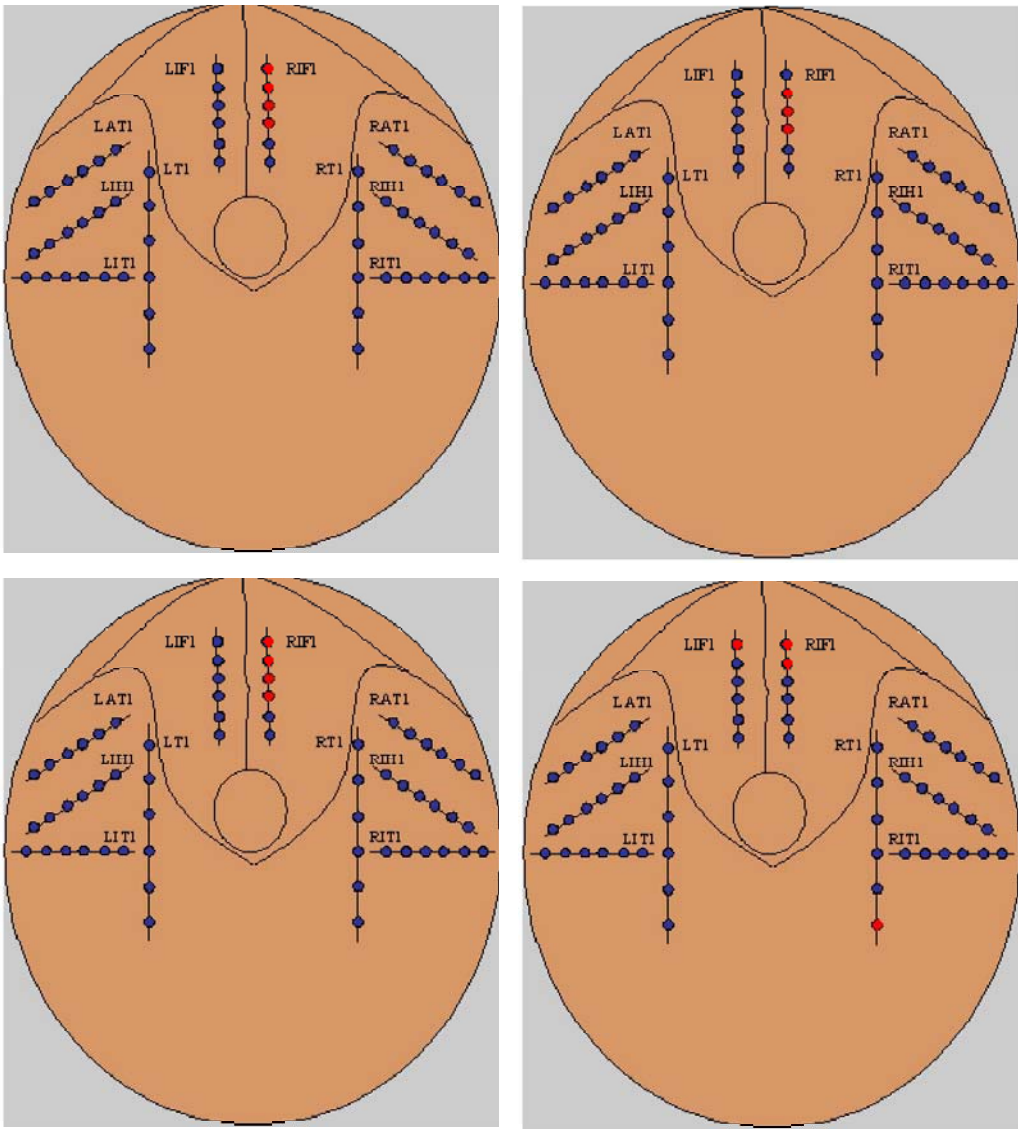


Figure 4.3-13 This figure illustrates the clustered electrodes for four different records from E05 at a support of 25%. The EHFO's were most consistently concentrated in electrodes from the RIF area, a portion of the SOZ.

4.3.5 EMRY07 (E07)

Ten interictal records were processed for this patient. Seizure onsets were apparent in electrodes LT1-4, LIT1-3, LIF1-2, and RIF1-4. The epileptologists who reviewed this record noted that the epileptic focus varied depending on the record (i.e., RIF, LIT, LT), making it difficult to localize. Epileptic oscillations were detected in several electrodes per record, however the overall results of applying FIM to the records returned consistent clusters of EHFO's within the SOZ, specifically LIT1-4.

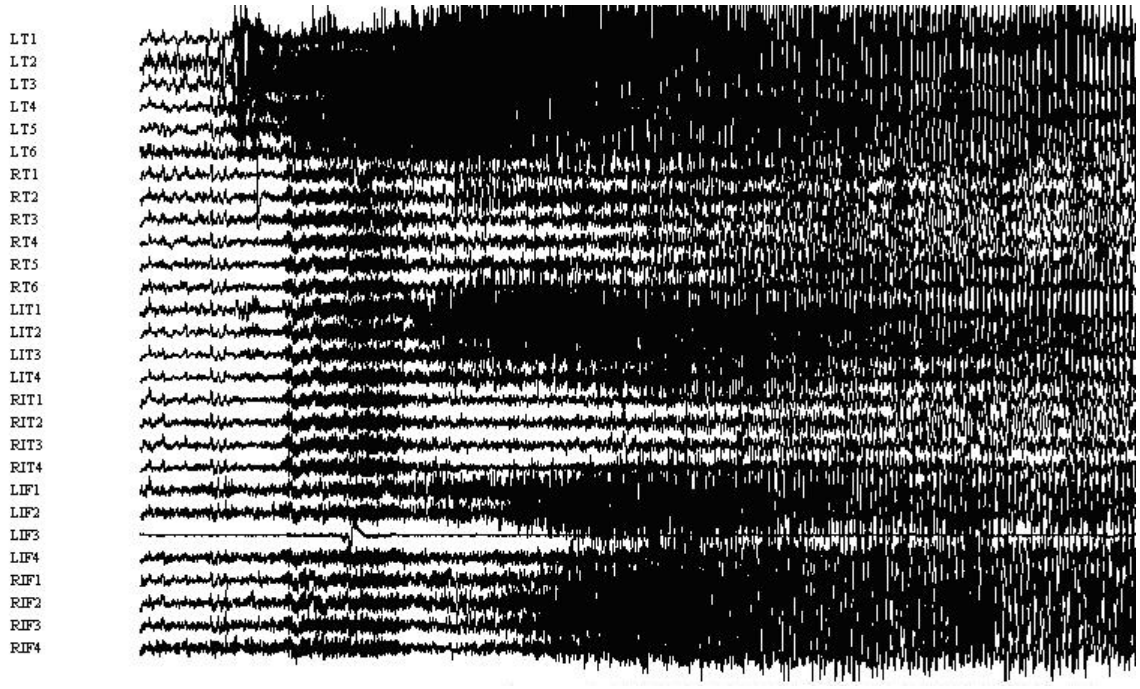


Figure 4.3-14 This figure illustrates the SOZ for E07, which variably included of electrodes in the LT, LIT, LIF, and RIF regions of the brain.

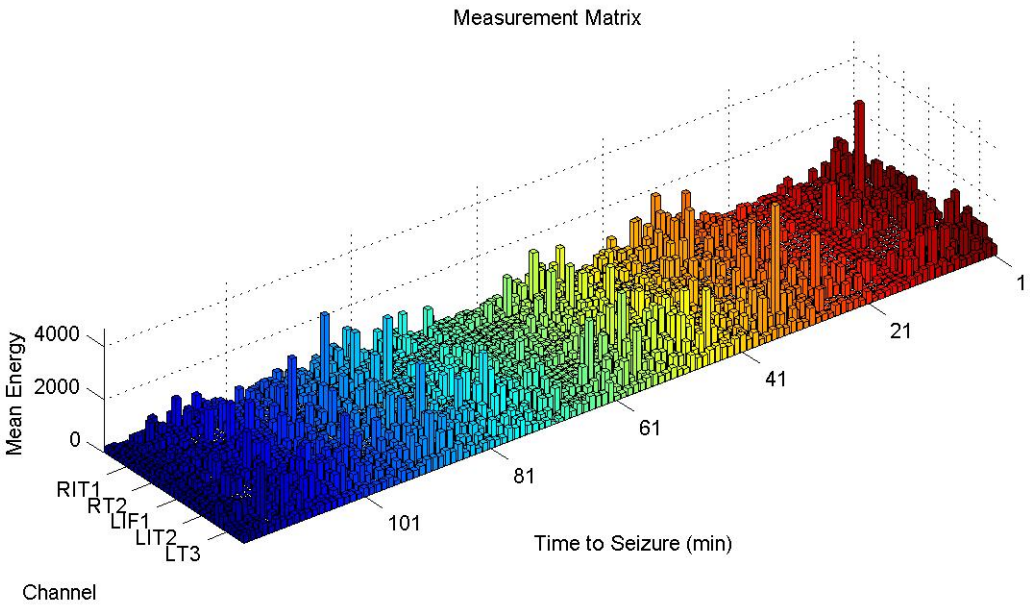


Figure 4.3-15 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject E07. It is the measure matrix for the interictal record captured in Figure 4.3-14.

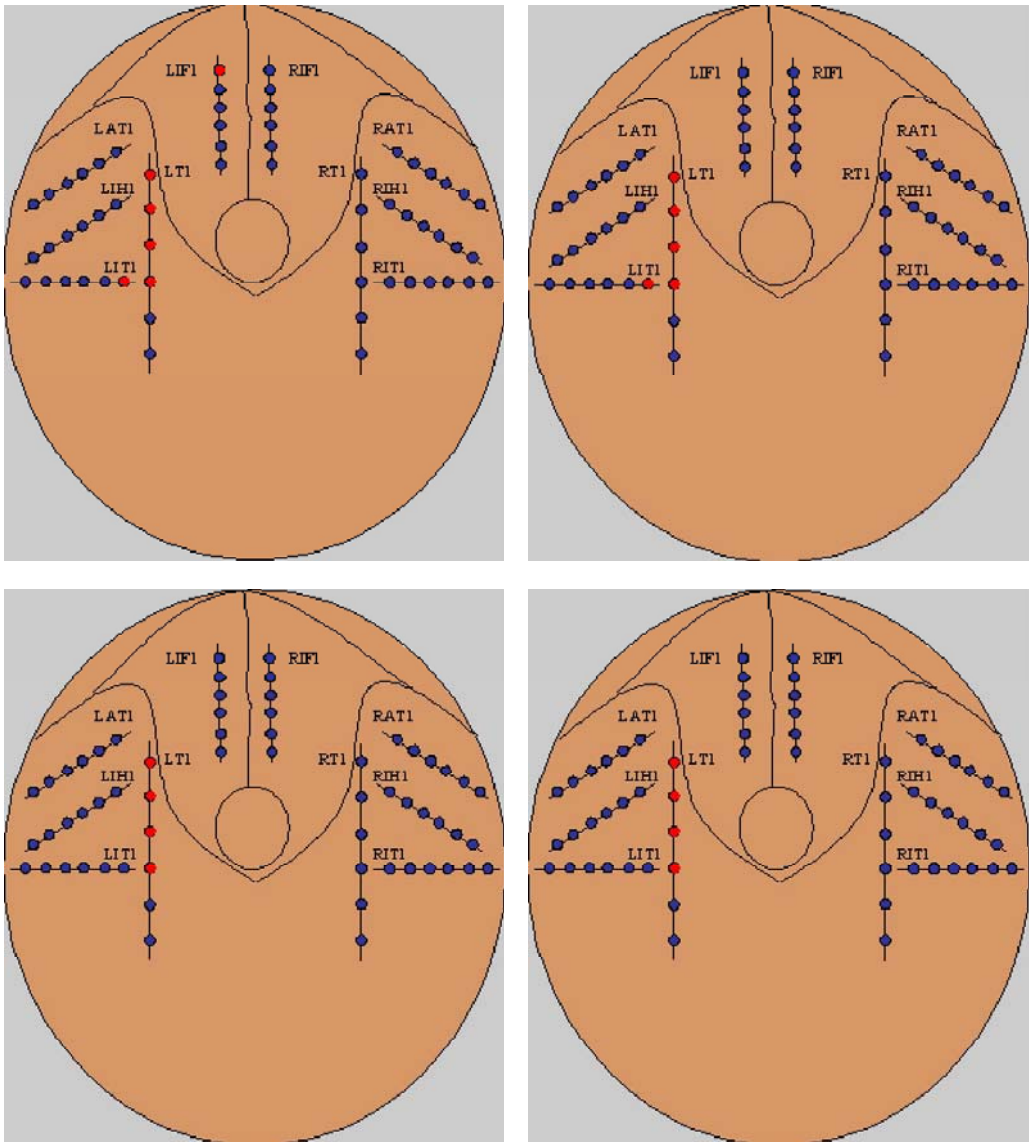


Figure 4.3-16 This figure illustrates the clustered electrodes for four different records from E07 at a support of 25%. Despite difficulty in localizing the epileptic focus by an epileptologist, the FIM consistently pinpointed electrodes in the LT area as an epileptogenic region according to clusters of high-energy EHFO's.

4.3.6 EMRY09 (E09)

Ten interictal records were processed for this patient. The epileptologists who reviewed this record observed a broad epileptic focus in the right area of the brain encompassing the regions RIF, RIT, and RT. However, seizure onsets transpired in nearly all the electrodes—left and right regions of the brain—after surveying all records. Epileptic oscillations were detected in several electrodes per record, and the overall results of applying FIM to the records returned consistent combinations of the electrodes within LT, LIT, RT, and RIT.

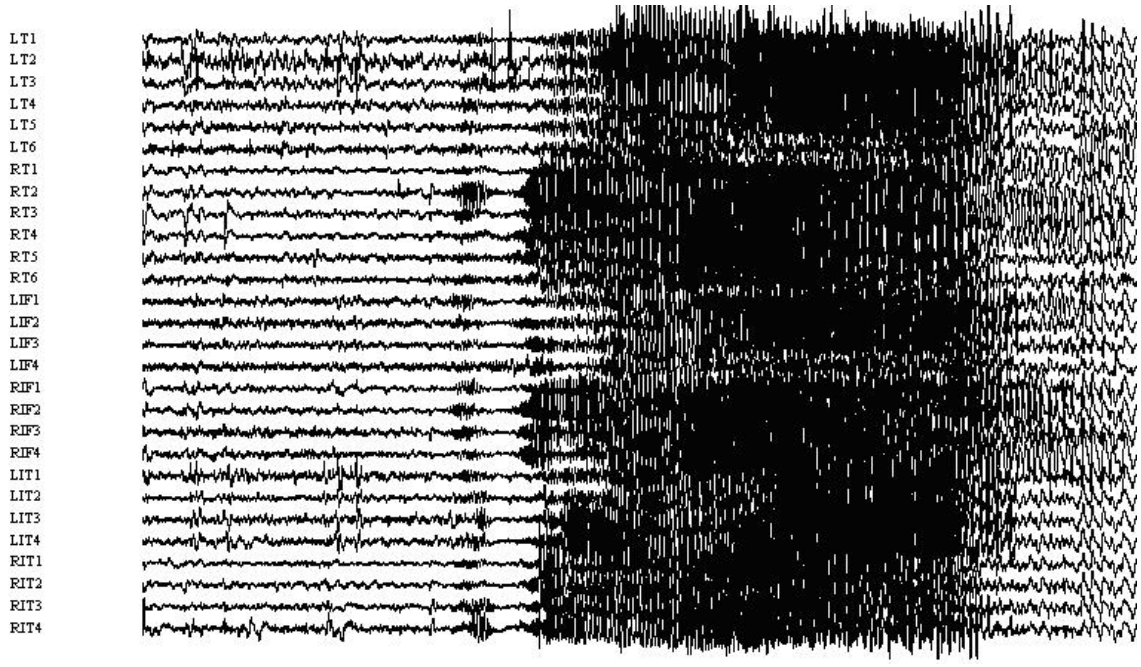


Figure 4.3-17 This figure illustrates the very broad SOZ for E09, where a seizure occurred in every region (e.g., LT, RIF) of the brain across several records.

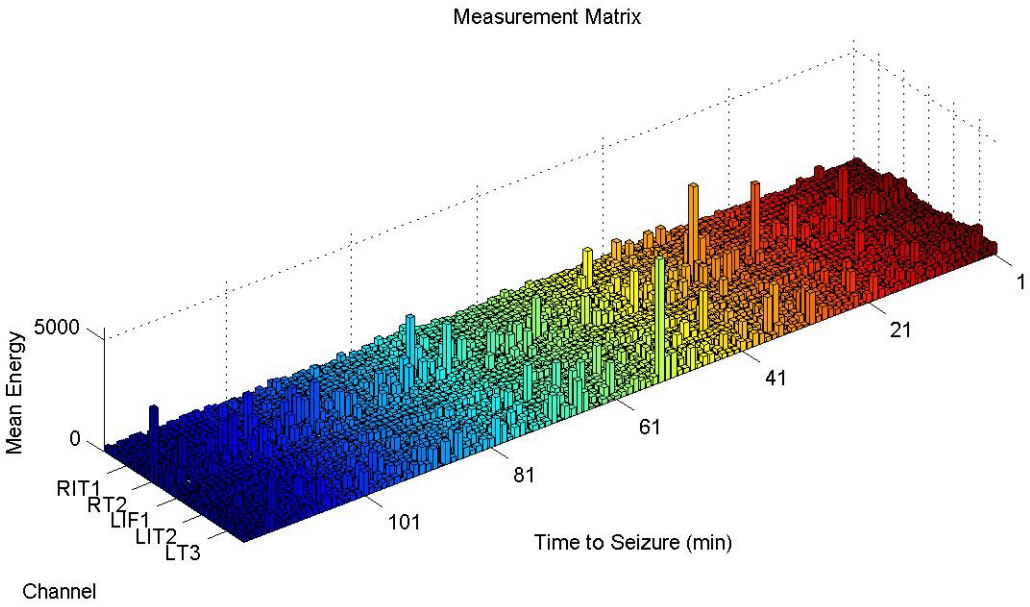


Figure 4.3-18 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject E09. It is the measure matrix for the interictal record captured in Figure 4.3-17.

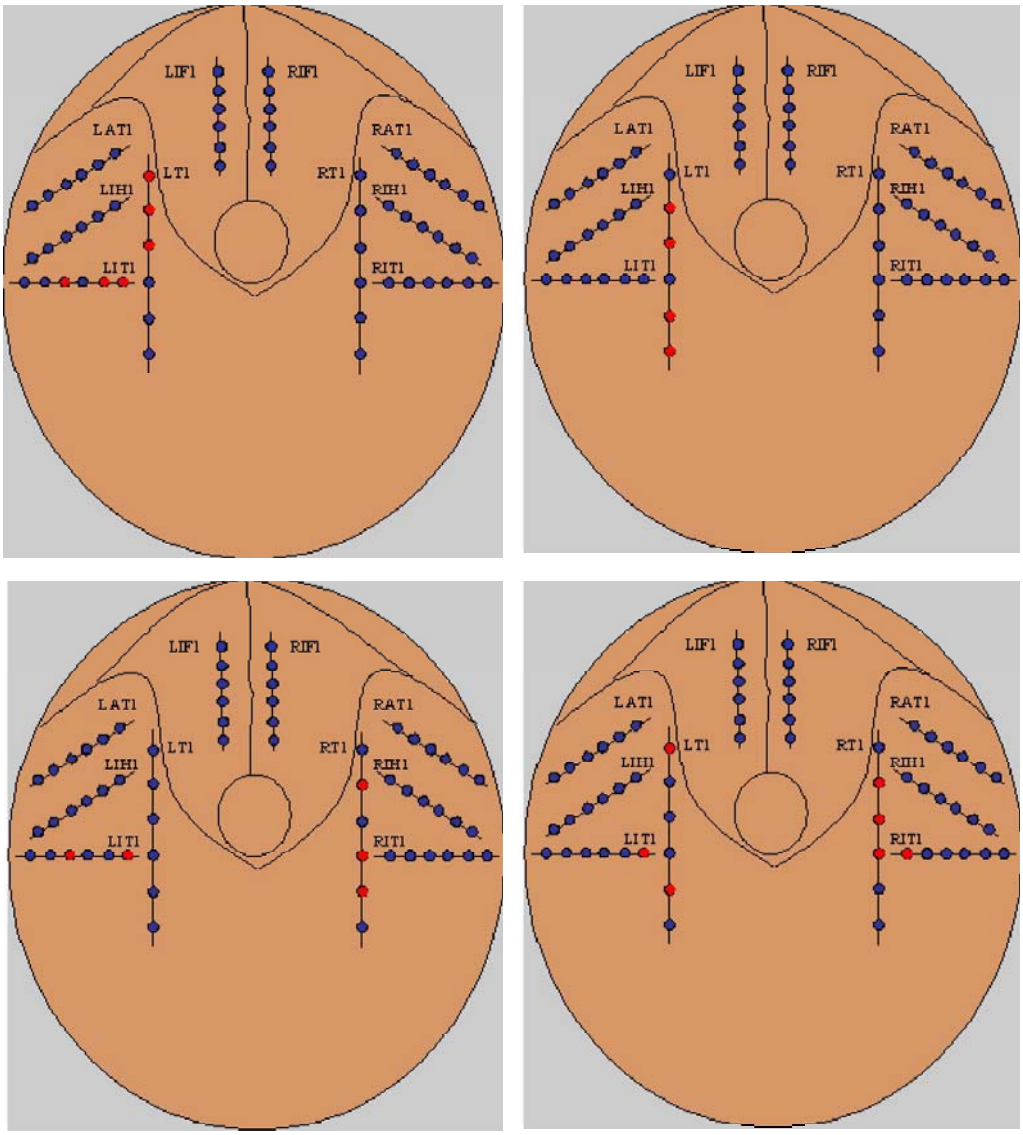


Figure 4.3-19 This figure illustrates the clustered electrodes for four different records from E09 at a level of support equal to 10%. In most records, the coupled channels included electrodes in the regions LT, RT, LIT, and RIT.

4.3.7 EMRY11 (E11)

Five records were available for this subject with the SOZ apparently localized to the right side of the brain in areas RIT, RT, and RIF. As displayed in Figure 4.3.7-1, measuring the putative epileptic events collected by the binary detector lead to comparable values of energy throughout nearly all the intervals and channels in each record. Because of this low discrimination in the measure matrices, mining the epileptic oscillations with such input would not reliably cluster electrodes. Figure 4.3-22 demonstrates the case where applying FIM would not provide clear decision-support.

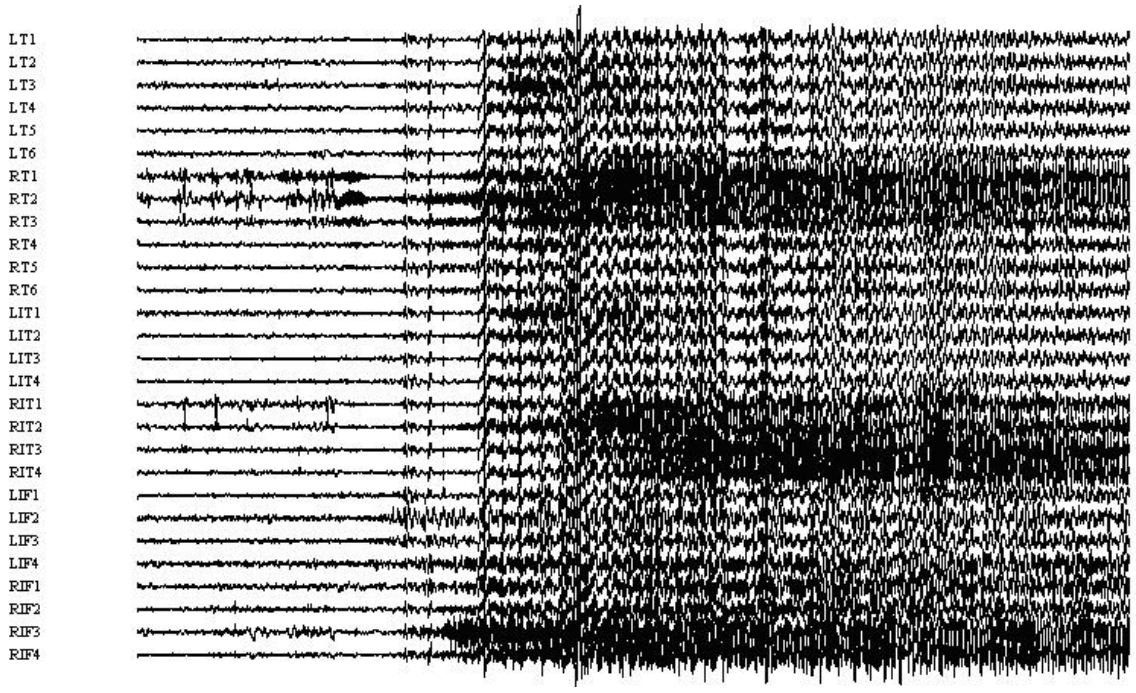


Figure 4.3-20 This figure illustrates the SOZ for E11, which included electrodes RIT2-4, RIF3-4, and RT1-3.

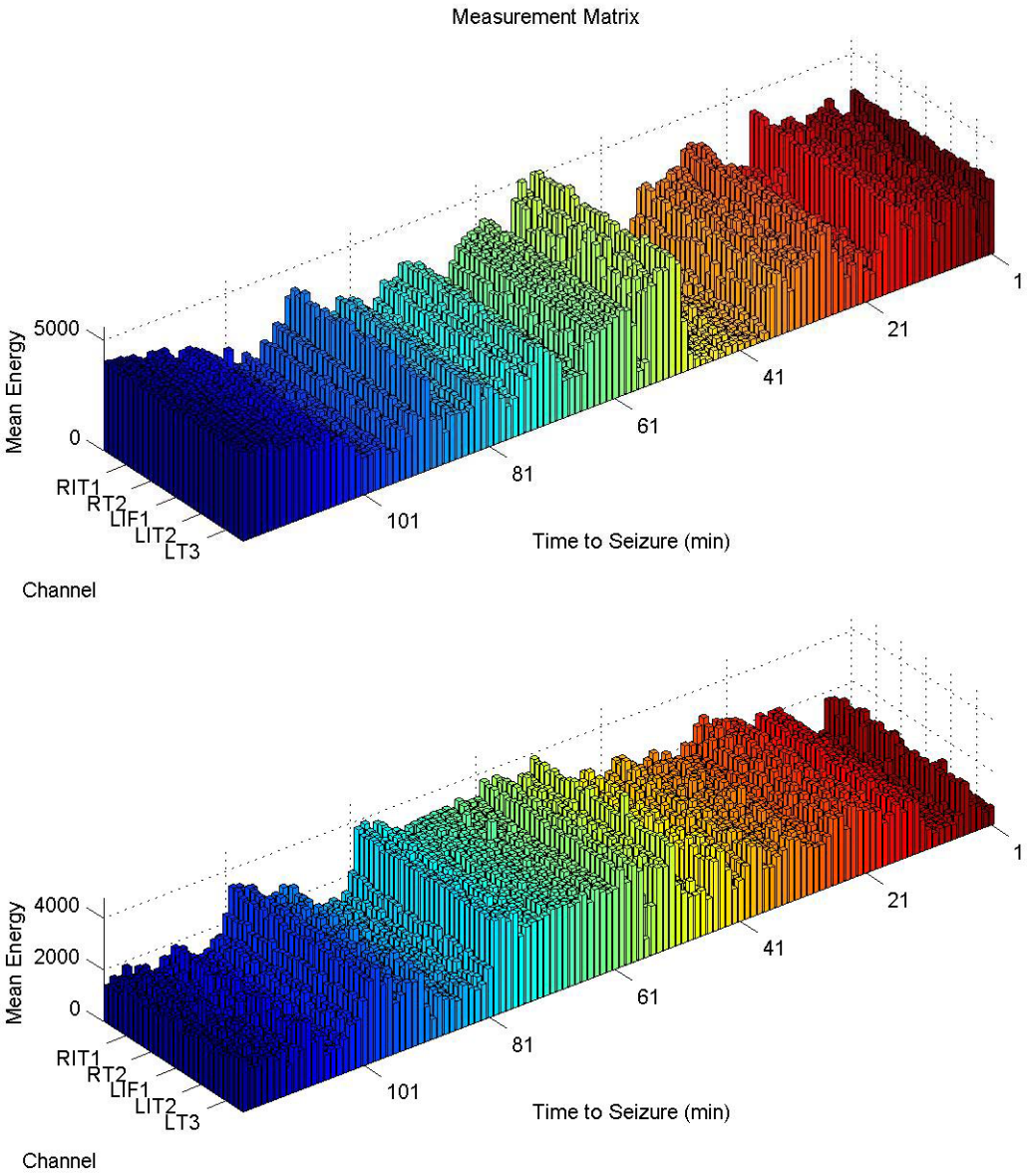


Figure 4.3-21 This figure exemplifies the input to the module that statistically clusters the spatial-temporal distribution of EHFO's using FIM for subject E11. The measure matrices for two different interictal records are displayed.

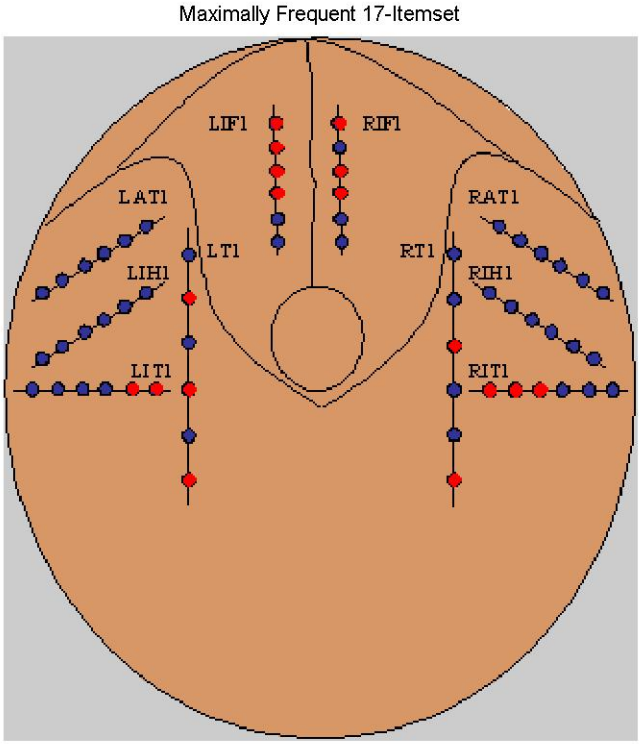


Figure 4.3-22 This figure illustrates the results of FIM applied to a measure matrix with poor discrimination of energy in detected interictal epileptic oscillations throughout the intervals and channels of a record from E11. The FIM was applied at a level of support equal to 69%. The apparent SOZ for this subject according to a physician was RIT, RT, and RIF. In most records for this patient, FIM could not putatively identify (couple) electrodes in that region as the epileptic zone.

4.3.8 Summary

In six of the seven subjects, applying FIM to multi-channel, long-duration interictal records in which epileptic high frequency oscillations were detected consistently yielded clusters of electrodes that typically formed regions near or within the seizure onset zone. Thus for a total of 52 out of 57 interictal records, FIM coupled electrodes that were primarily located within the SOZ, supporting hypotheses that EHFO's might be useful in pinpointing the SOZ. Furthermore, in the occasions when interictal oscillations were clustered in channels outside of the SOZ, the FIM likely identified neurodegenerative tissue in the epileptogenic zone since the events were often associated with ictal activity in the SOZ.

Reviewing the IEEG of the only patient for which using FIM would not have been successful, the following sources of error were considered: 1) the patient possessed training data with high SNR but the testing data exhibited poor SNR, resulting in the creation of a feature with high quality (separation between epileptic and non-epileptic interictal oscillation) that was ill-conditioned for proper detection; and 2) the a-feature was correct while the detections appearing spurious were true epileptic oscillations that were extremely minute and indistinct. If the former case were true, a set of training data that included the noise from the testing data—providing a more adequate sample for crafting the artificial feature and classifying the testing data—would have helped.

4.4 Characterization of Epileptic High Frequency Oscillations

Having designed a means to accurately detect and mine the epileptic high frequency oscillations, this section was dedicated to characterizing the events. The following techniques of analysis were considered:

- Regression of normalized histograms for specific measures (i.e., *amplitude*, *frequency*, *phase, duration*) of the aberrant oscillations.
- Hidden Markov Modeling using regressed probability mass functions for measurements of the epileptic events.
- Reproducibility of the selected genetically programmed c-features. That is observing whether or not the GP consistently selected certain classical features that were input for the stochastic global search algorithm.

4.4.1 Regression of Normalized Histograms

The seemingly sinusoidal behavior of the epileptiforms lead to the presumption that the oscillations could be described as a random process with the equation $ehfo(t) = \alpha \cdot \cos(2\pi\varphi t + \theta) + \varepsilon(t),\ 0 \le t \le \tau$, where the random variables $\alpha, \varphi, \theta, \tau,$ and $\varepsilon(t)$ were parameters for the amplitude, frequency, phase, duration, and additive noise for the signal. The aim was to potentially make sense of the epileptic oscillations by regressing the probability mass functions for $\alpha, \varphi, \theta,$ and τ since probability functions convey important information about a variable—or the process yielding a variable—in the theory or application of a physical problem. In addition, the events could be modeled (Section 4.4.2) using statistics of the regressed probability density functions.

The probability distribution function for $\varepsilon(t)$ was assumed to be a standard normal distribution. Figure 4.4-1 illustrates the probability mass functions that characterize each measure of the abnormal electrographic activity. For each sample of measures, several distributions (e.g., Normal, Log-Normal, Weibull, Rician) were regressed and the log-likelihood ratio in Equation 4.4-1 was computed for each regression.

$$\Lambda(x) = \frac{\sup\{L(\eta \mid x) : \eta \in \mathrm{H}_0\}}{\sup\{L(\eta \mid x) : \eta \in \mathrm{H}\}} \tag{4.4-1}$$

In Equation 4.4-1, considering a single distribution that was regressed on a sample of measures, *x*, the numerator of the likelihood-ratio was the maximum of the likelihood function, $L(\eta \mid x)$, under the null hypothesis that parameters of the distribution, η, equal specific values, H_0, and the denominator was the maximum likelihood when the parameters can take on many values in the parameter space, H. The distribution with the highest log-likelihood ratio, Λ, determined the best regression. Alternatively, the distribution with the lowest magnitude of the log-likelihood ratio, $|\Lambda|$, determined the best fitting curve for the observed sample.

Interestingly, the best regression for $\alpha, \varphi,$ and τ resulted in the same probability curve, or sum of such distributions in the case of the frequency of the events: an *inverse Gaussian distribution*. The probability density function for the inverse Gaussian is $pdf(x) = \sqrt{\frac{\lambda}{2\pi x^3}} e^{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}}$; where the mean of the PDF is μ and the standard deviation is

$\frac{\mu^3}{\lambda}$ for the random variable *x*. In contrast, the best fitting distribution for the phase of the measured epileptic oscillations was the uniform distribution, which was expected recalling typical problems to define wide sense stationary random processes in the theory and literature for communications systems.

Figure 4.4-2 captures the trend in log-likelihood across all feasible distributions and all the measures excluding the phase-angle. Meanwhile, Figures 4.4.1-3 through 4.4.1-5 are the results from bootstrapping the mean measures of the epileptic oscillations and computing the mean for each bootstrapped sample to provide perspective on how the mean and spread for each measure of an EHFO under the inverse Gaussian distribution varied per subject. For each subject, bootstrapping the mean value of the phase yielded coincident probability distributions centered at $\frac{\pi}{2}$ as represented in Figure 4.4-6.

Considering that results of the stochastic regression lead to an inverse Gaussian distribution in characterizing parameters of an abnormal action potential preceding seizures and that an inverse Gaussian distribution describes the distribution of the time taken by a particle in *Brownian motion* to reach a fixed distance [172, 173], a classical model for the firing behavior of a simple network of neurons came to mind [170, 171] as a plausible perspective on the epileptic high frequency oscillations in the EEG. More specifically, the work of Gerstein and Mandelbrot [170] proposed the following paradigm, which satisfactorily suited observed data and simulations via Monte Carlo methods: if the electrical potential of a neuron could be quantified while excitatory (inhibitory) inputs increase (decrease) the membrane to (from) an action potential, or firing potential, then the time the membrane fires can be approximated by the *first hitting*

time of a certain threshold of electrical and chemical synapses for a Brownian motion with drift. The first hitting time, or *hitting time*, is simply the first time in which a random process reaches an inevitable stopping condition before the process is renewed.

Briefly, Brownian motion refers to the physical phenomenon of the random movement of particles suspended in a fluid due to the instantaneous imbalance in the force exerted by the small liquid molecules on the particle. An analogy that can help understand this concept is the movement of a large balloon—suppose 10 meters in diameter—through an area widely crowded with people who hit the balloon at different times and in different directions with the motions being completely random [172] after an initial period of rest. At a certain time, depending on the number of persons hitting the balloon, the force (direction and magnitude) at which they hit the balloon, the balloon may move due to an imbalance forces (the sum of the forces is not zero). If such imbalance exists at all times, the people generate random motion for the balloon. Similarly, since a water molecule is about 1 nm and the pollen particle is roughly 1 μm in diameter—approximately 1000 times larger than a water molecule—the pollen particle can be considered as an extremely large balloon that is constantly being pushed by water molecules [172].

Combining the abovementioned notions, this research supposed that the action potential for an epileptic oscillation resulted after a network of dysfunctional neurons in synaptic transmission (movement) until a cluster of degenerate neurons (Brownian particles) fired, having reached the firing threshold (hitting time) for the network, before repeating the process. Moreover, this work hypothesized that the inverse Gaussian probability density function aptly described the three measurable parameters directly

related to how the neuronal cluster attained sufficient electro-chemical potential (through superposition) to yield an action potential—essentially a noisy sinusoid: amplitude, frequency, and duration. Furthermore, the voltage recorded by an IEEG electrode (the balloon) during an epileptic oscillation perhaps measures the action potential (total force) of many dysfunctional neurons (particles) rapidly firing (force of hitting the balloon) for a period of time (force duration) near the electrode at different times (force frequency) with different amplitudes (force magnitudes).

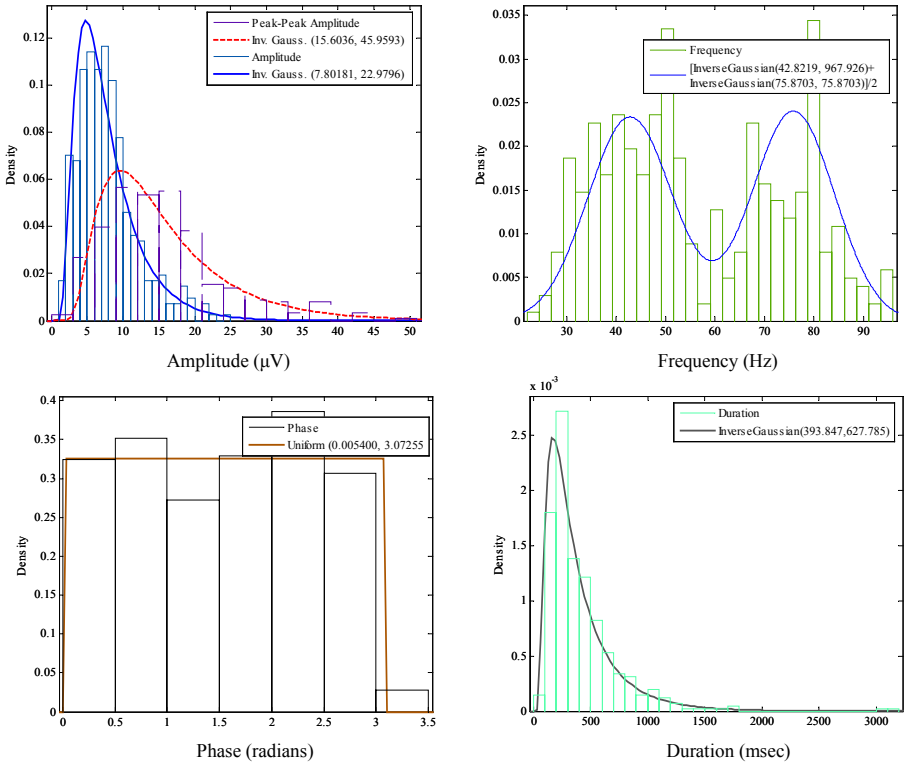


Figure 4.4-1 This figure illustrates the results of regressing the histograms for the amplitude and peak-to-peak-amplitude (*top left*), frequency (*top right*), phase (*bottom left*), and duration (*bottom right*) for a sample of epileptic oscillations. The Inverse Gaussian distribution best fit all measures except the phase, for which a uniform distribution on the interval 0 to pi was apt.

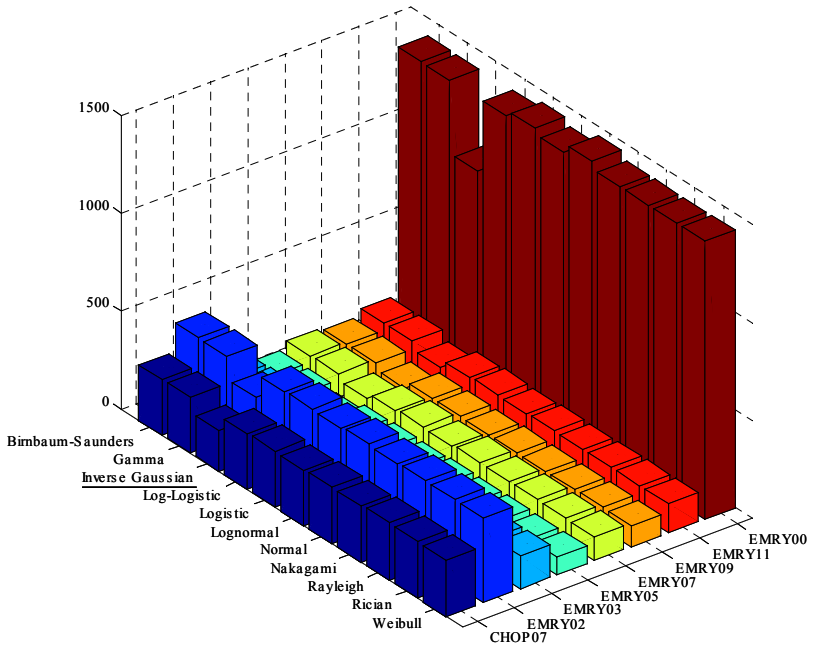


Figure 4.4-2 This figure represents the finding that the Inverse Gaussian distribution (*underlining*) best fit all measures, excluding the phase, according to the magnitude of the log-likelihood ratio (*vertical axis*). The results are illustrated for measures that were stratified by subject and aggregated for all subjects (EMRY00), revealing a unique circumstance in which the Inverse Gaussian distribution exhibited the lowest values of the negative log-likelihood ratio in each subject as well as all combined subjects.

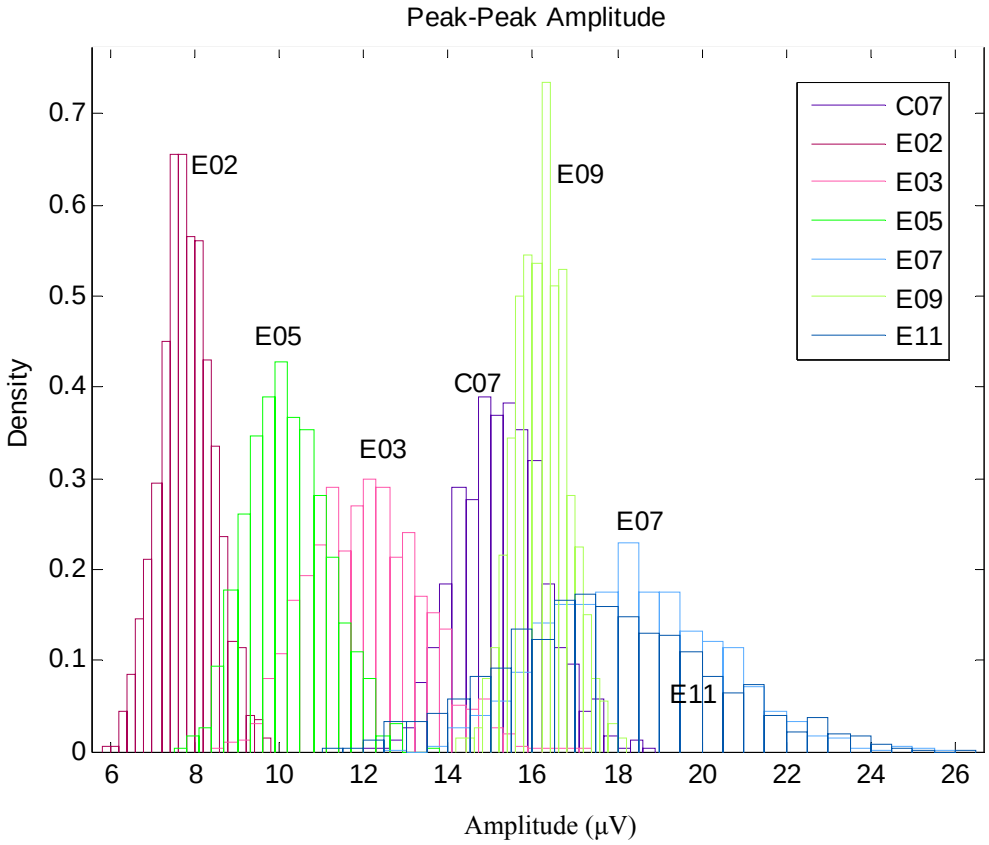


Figure 4.4-3 This figure captures the mean and the standard deviation of the Inverse Gaussian distribution characterizing the peak-to-peak amplitude of an epileptic oscillation for each of the seven subjects.

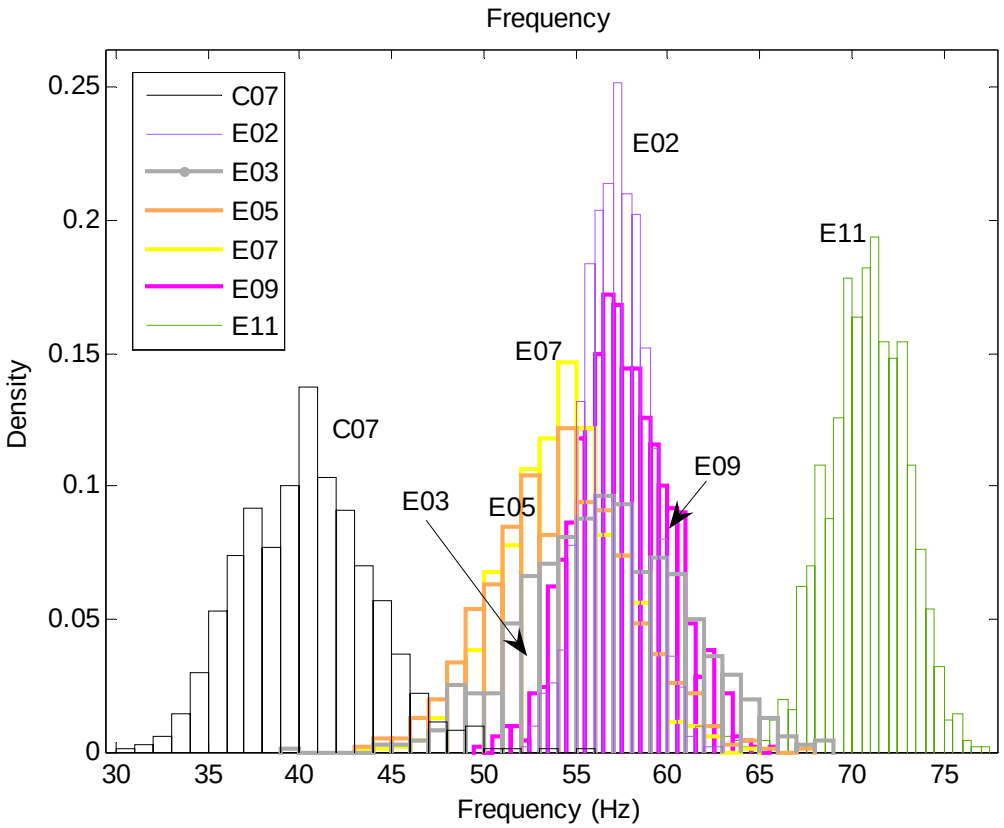


Figure 4.4-4 This figure captures the mean and the standard deviation of the Inverse Gaussian distribution characterizing the frequency of an epileptic oscillation for each of the seven subjects.

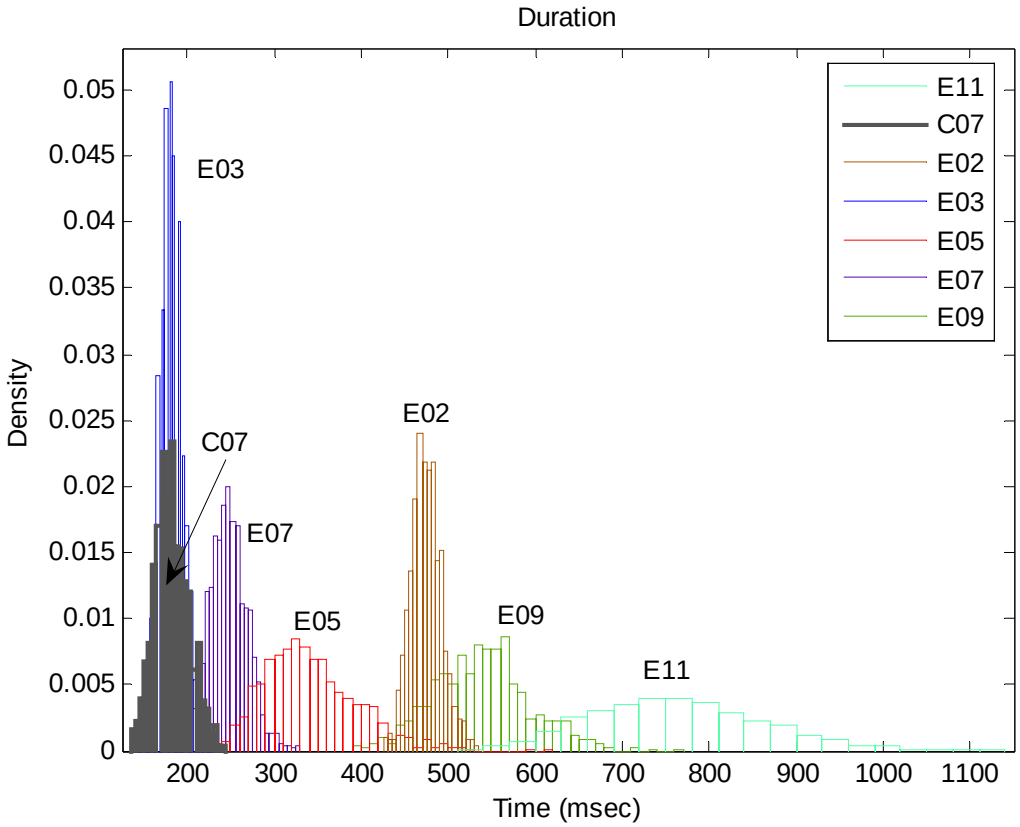


Figure 4.4-5 This figure captures the mean and the standard deviation of the Inverse Gaussian distribution characterizing the duration of an epileptic oscillation for each of the seven subjects.

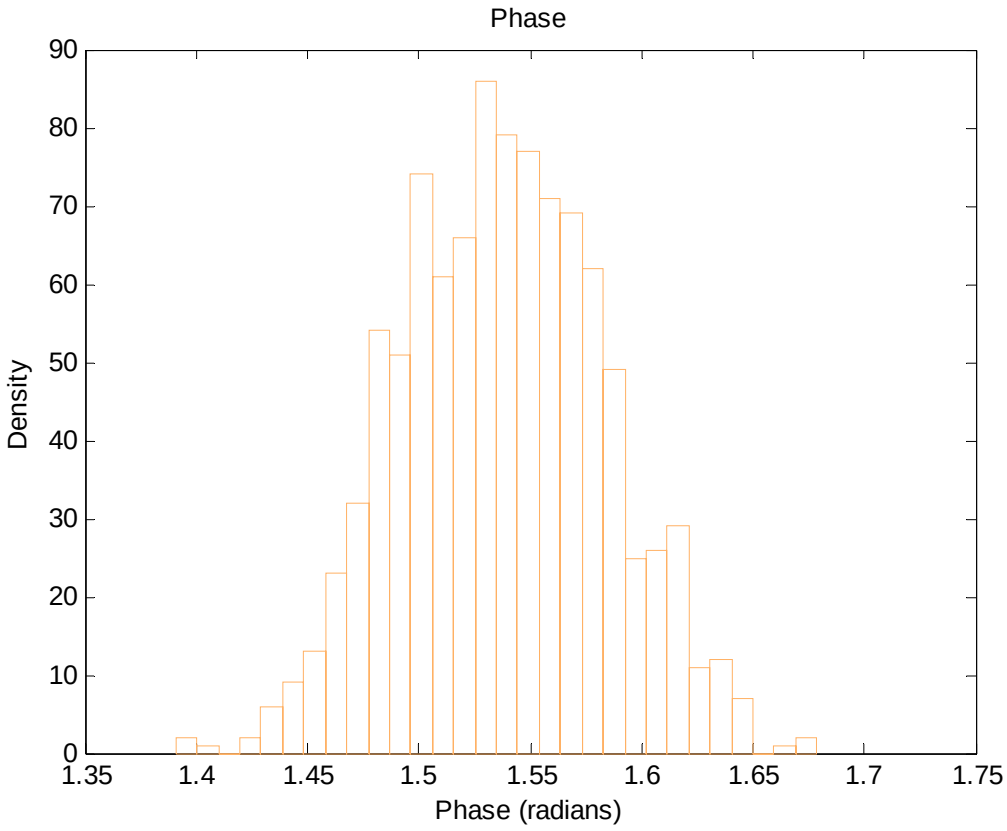


Figure 4.4-6 This figure captures the mean and the standard deviation of the Uniform distribution characterizing the phase of an epileptic oscillation for each of the seven subjects. A single PDF is illustrated to point out that each was practically the same and all centered near $\frac{\pi}{2}$.

4.4.2 ***Hidden Markov Modeling***

Regressing the histograms not only provided a possible analogue as for understanding the physics of the epileptic oscillations preceding seizures but also supported a simple model to simulate an EEG signal with the pathological events examined in this research. Although it was not intended to develop a system that rigorously reproduced interictal EEG, expanding such a model to incorporate other normal and abnormal electrographic waveforms (e.g., seizures, theta rhythms, spikes) could be useful in optimizing certain techniques (e.g., electrical stimulation) to effectively treat epilepsy, especially when obtaining actual data from humans for that purpose is sparse or unethical. In addition, the model could readily facilitate a proof of concept for proposed methods to analyze EEG from subjects with epilepsy. Furthermore, the model could eliminate the burden to mark epileptic oscillations for training classifiers and algorithms to produce optimal features.

A discrete time *hidden Markov model* (HMM) proved an adequate framework to replicate electroencephalograms from a patient with epilepsy. In a discrete time HMM, a sequence of unobservable states at successive times transpires in which each state emits a visible symbol. A *transition probability matrix*, or *transition matrix*, contains likelihoods of leaving one state for another, while an *emission probability matrix* or *emission matrix*, contains likelihoods of returning an output in the current state. A HMM matched the application of generating synthetic data because one should observe a particular type of electrographic activity but not the state of the model that dictated the artificial output. In fact the HMM was apt considering the review of actual EEG. References [166-168] can be reviewed for more detailed information on HMM.

The simple HMM in Figure 4.4.2-1 controlled the addition to a filtered, random white-noise process of an epileptic event (state 3) according to $ehfo(t) = \alpha \cdot \cos(2\pi\varphi t + \theta),\ 0 \le t \le \tau$, an oscillation epitomizing the mixture of an alpha rhythm and a beta rhythm (state 2) according to $v(t) = \alpha_1 \cdot \cos(2\pi\varphi_1 t + \theta_1) + \alpha_2 \cdot \cos(2\pi\varphi_2 t + \theta_2),\ 0 \le t \le \tau_{ab}$, or no event for a duration of one second (state 1), thereby modifying the efforts of Penny and Roberts [169]. The parameters $\alpha, \varphi, \theta, \text{and}\, \tau$ in $ehfo(t)$ were determined by the regressed probability density functions and corresponding statistics in Section 4.4.1, while the parameters for the amplitude (μV/mm), frequency (Hz) and phase of the mixture of rhythms, *v(t),* were presumed to possess the following probability density functions in that order:

- Uniform (a=0, b=1) for τ_{ab}, α_1 *and* α_2.
- Uniform (a=8, b=12) for φ_1 and Uniform (a=12, b=20) for φ_2.
- Uniform (a=$-\pi$, b=π) for θ_1 *and* θ_2.

Observing the real interictal IEEG, arbitrary values for the time-independent probabilities of the transition matrix (*A*) were set. Figure 4.4.2-1 and Figure 4.4.2-2 demonstrate the interictal IEEG generated by a model with $A = \begin{bmatrix} .80 & .15 & .05 \\ .80 & .20 & .00 \\ .95 & .00 & .05 \end{bmatrix}$. The emission matrix equaled the identity matrix, or $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ since the emissions were considered two mutually exclusive events.

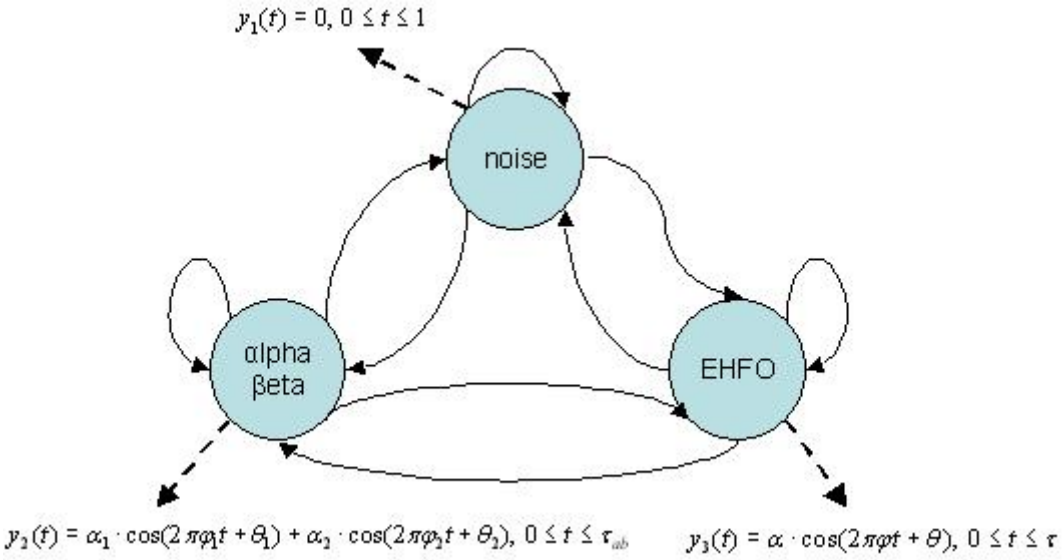


Figure 4.4-6 This figure illustrates the basic HMM used to simulate interictal IEEG.

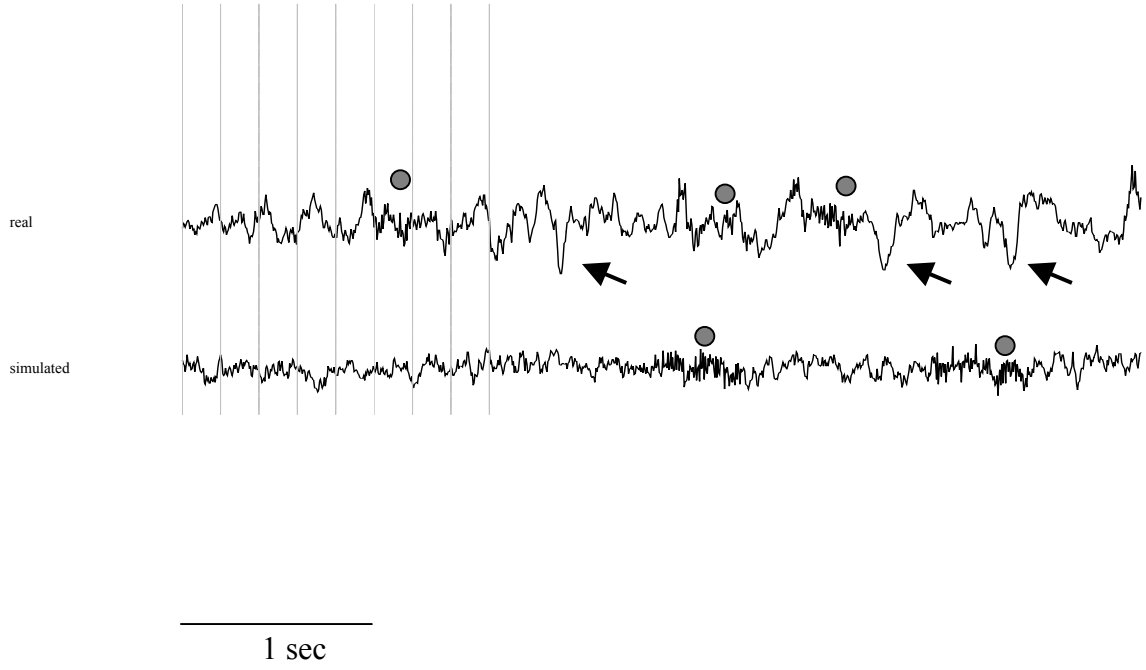


Figure 4.4-7 This figure juxtaposes data from a real EEG signal (*top*) and a synthetic EEG signal (*bottom*) that was produced via a hidden Markov model with arbitrarily defined parameters. The signals are very similar in content, including the epileptic events (dots). Improvements to the simulated data can include states that generate the slow waves (arrows) in the real EEG as well as other typical signals excluded from the HMM.

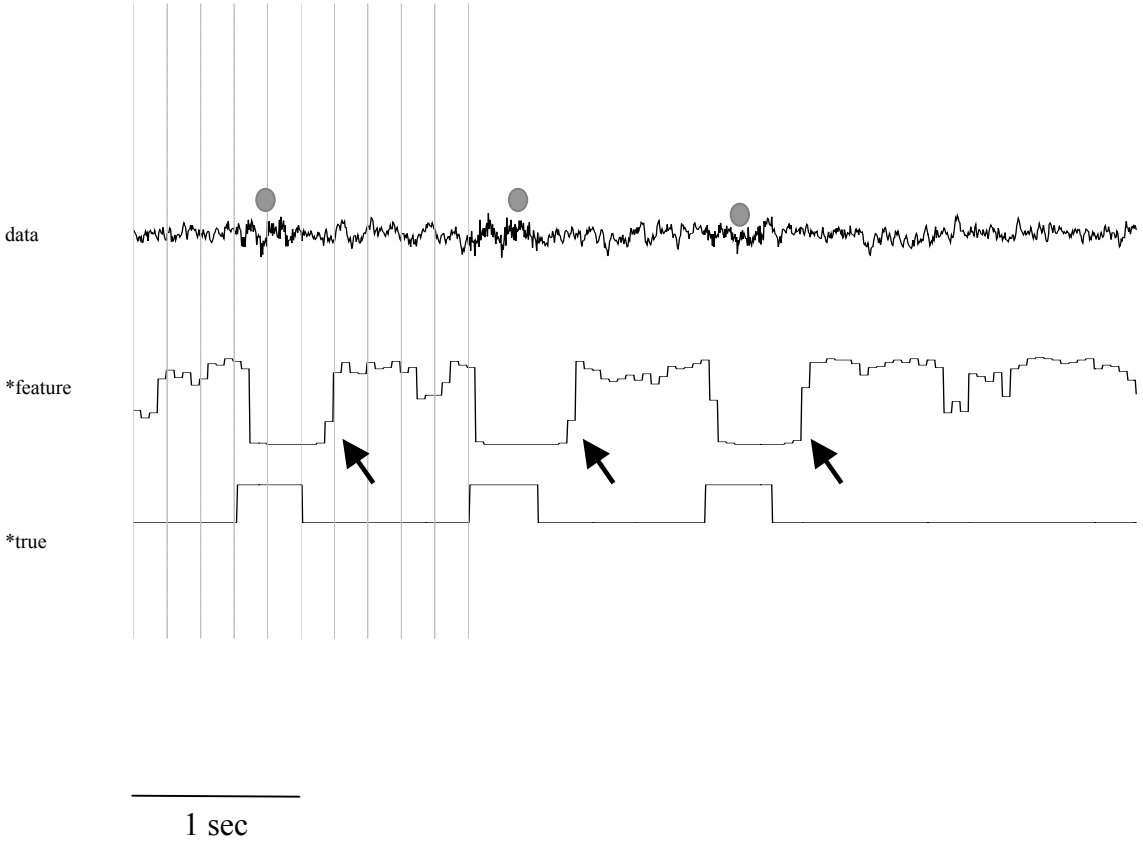


Figure 4.4-8 This figure illustrates the success of the model to make an artificial EEG signal (*top*). When the synthetic data was processed with a genetically programmed feature (*middle*) created from the epileptic high frequency oscillations and normal activity in actual electrographic recordings, the resulting feature-series highlighted (*arrows*) the epileptic events (dots) that the HMM generated according to the corresponding hidden state (*bottom*).

4.4.3 ***Reproducibility of the C-Features Selected by the Genetic Programming***

When the GP initializes, a random population of solutions for feature-fusion (selection and transformation of c-features) is generated. This first population certainly affects the final population of possible solutions from which the optimal function for fusion is chosen, but is the effect significant; and more importantly, does the final function consistently choose a set of features or feature-domains? If so then certain features and/or feature domains clearly distinguish an abnormal event, such as an EHFO, from normal baseline activity thus characterizing that event.

This question was investigated by setting up the query as a frequent itemset mining problem, where the item was a feature, making all itemsets a subset of features from the complete list of features given to the GP. The GP was run 100 times, at varying number of generations (i.e. 5, 10, 20, 30, 35) per trial for each subject. Once the 100 solutions per subject were obtained, each solution was parsed to extract the subset of features chosen by the GP to create an optimal feature. Consider now the 100 trials for a specific number of generations for a particular subject. The 100 subsets of features resulting from the trials were input to the frequent itemset mining algorithm, which with statistical rigor returned the most frequent subset of features contributing to a genetically programmed feature. Thus, if the FIM returned set A containing, for instance, the features curve length, entropy, and kurtosis as the maximally frequent itemset for 53 occurrences, then more than half of the time that the GP is run it will likely produce an optimal artificial feature relying on at least those three features. Such a result would not only demonstrate some consistency in the heuristics of the GP to pick features but also suggest something interesting and important about the few features that were selected.

Below Table 4.4.3-1 and Table 4.4.3-2 portray the propensity of the genetic programming algorithm to pick the same subset of features. Table 4.4.3-1 contains the subsets represented by a row of indices (e.g., 6 5 10 1, 8 6 1) taken from the complete set of features—taken from Appendix B—listed in Table 4.4.3-2 with the number of occurrences for a particular subset in parentheses and the most frequent item in the itemset (subset of features) as the last index in the row. It was interesting to note that the GP consistently *never* returned a set with only one feature, although curve length often appeared in the majority of subsets. Furthermore, the root mean square did not appear as frequently in subsets as did curve length or peak frequency, meaning that optimal features to distinguish epileptic and non-epileptic oscillations were routinely achieved without any needed contribution from RMS.

In most subjects, the number of generations of the GP affected the subset that FIM returned. For patients E02 and E11, however, this was not the case because the FIM always clustered the features peak frequency, mean periodogram magnitude, kurtosis, and curve length for E02 and the features peak frequency, mean periodogram magnitude, and curve length for E11. For the remaining subjects, the at least one of the features in the maximally frequent itemset for a specific number of generations of the GP remained in the maximally frequent itemset for each of the other number of generations, again implying worth in the particular feature over others because of its recurrence. For example, the max frequent itemset for patient E09 varied in size and items across, yet consistently the GP selected peak frequency, which was mined in every itemset. A similar finding held for subject C07 and curve length. In sum, these results evince the notion that one precise classical feature (or set of c-features) is not apt for all subjects.

Table 4.4-1 List of classical features inputted to the GP for the experiment to evaluate the reproducibility of selected subsets of features.

	Generations				
Subject	5	10	20	30	35
C07	4 1 (18)	3 1 (25)	3 1 (25)	3 1 (26)	2 3 1 (10)
E02	6 5 10 1 (17)	6 5 10 1 (24)	6 10 5 1 (26)	6 5 10 1 (33)	6 5 10 1 (30)
E03	11 1 (15)	6 1 (18)	6 1 (15)	11 1 (10)	11 1 (17)
E05	1 2 (15)	1 4 (15)	4 1 (20)	6 10 2 (10)	10 2 (16)
E07	1 4 (17)	8 4 (16)	2 4 (15)	10 4 (16)	1 2 (16)
E09	4 6 2 (10); 6 2 1 (10)	4 2 6 (13); 2 6 1 (13)	8 6 1 (13)	8 6 4 2 (11)	2 6 1 (13)
E11	6 5 1 (17)	6 5 1 (19)	6 5 1 (14)	6 5 1 (16)	6 5 1 (22)

Table 4.4-2 List of classical features inputted to the GP for the experiment to evaluate the reproducibility of selected subsets of features.

01	**CurveLength**
02	**RootMeanSquare**
03	**MeanRectifiedValue**
04	**Crossings**
05	**MeanPeriodo**
06	**PeakFreq**
07	**Iqr**
08	**Std**
09	**Skewness**
10	**Kurtosis**
11	**ShannonEntropy**
12	**RenyiEntropy**
13	**Complexity**

Chapter 5: Conclusion & Future Research

5.1 Conclusion

This work advances the state of the art in research on epilepsy by providing a sound, general framework for the automatic detection and analysis of epileptic oscillations given multi-electrode, intracranial electroencephalograms recorded for long durations prior to ictal onset. Appropriately, this framework answers a substantial need demanded by convergent literature, scientific meetings, and medical discussion: a diagnostic tool to reliably map an epileptic network of dysfunctional brain represented by clusters of intracranial electrodes with frequent bursts of epileptic interictal activity. More importantly, this work provides decision-support for epileptologists to more effectively study and the treat millions of persons worldwide who are afflicted with epilepsy.

The presented approach to detect interictal epileptiforms bests a naïve "benchmark," achieving measures of performance with statistically significant differences between the benchmark and new methods over testing data from six subjects according to a two-factor ANOVA. The application of an evolutionary algorithm, chiefly genetic programming, to create an artificial feature that optimally distinguishes between epileptic and non-epileptic events classes is the essential reason for the improvement over the benchmark. The use of evolutionary algorithms readily resolves recurrent issues in previous work to develop an adequate detector of interictal epileptiforms, namely the scaling of features to prevent misclassification due to dominant (with larger variance) features, the selection of primary features for the best class-separation, and the fusion of

primary features considering the *curse of dimensionality*, which implies that high-dimensional feature-spaces have the potential to be very complex relative to low-dimensional feature-spaces.

Moreover, this work introduces the computation of an ROC curve to evaluate the quality of a feature for the detection of interictal epileptic oscillations, showing that high quality features exhibit high values for the area beneath the ROC and that evolutionary algorithms could produce a-features with higher quality than c-features. In fact, scatter-plots of AUC (projected performance) versus k-factor (class-separation) and regression of nonlinear curves within the plots revealed that evolutionary algorithms produce high-quality features, especially if the heuristics converge to high k-factors, a result that depends upon the SNR of the training data for each subject. Additionally, a detector using a *k*-nearest-neighbor rule for classification rather than the conventional choice of a threshold typically performs better the benchmark across subjects because the *k*-NN rule better classifies the nonlinear feature-space dictated by the data in this problem. Therefore, for a binary detector of epileptic oscillations, this work demonstrates that an evolutionary algorithm is an excellent technique to provide a feature and nonlinear classifiers, such as a *k*-NN, were more appropriate than simple thresholds.

Identifying epileptogenic regions, for instance the seizure onset zone, as a subset of electrodes among many intracranial electrodes is delineated as a mining problem. Applying a frequent itemset mining algorithm to multi-electrode, long-duration interictal data from six subjects after detecting epileptiforms within the records demonstrates that FIM can reliably map epileptic networks because the algorithm routinely outputs certain clusters of intracranial electrodes with high concentration of interictal epileptic energy,

which coincided with the seizure onset zone in each record. Because the epileptic interictal oscillations are primarily concentrated within the SOZ, this work not only supports hypotheses that epileptic oscillations might be useful in pinpointing the SOZ but also suggests that the FIM isolates sites of dysfunctional tissue when the events clustered in channels near (outside of) the SOZ.

Calculating histograms and regressing probability density functions for selected features of the epileptic oscillations in the training data of this study helps to characterize the epileptiforms as a multivariate random process, facilitate the use of a simple hidden Markov model to simulate interictal IEEG for the first time in literature, and explain the occurrence of the oscillations from damaged neurons with Brownian motion as a plausible physical analogue. Building upon this analysis by applying it to other waveforms in the IEEG may provide more depth or breadth in understanding the mechanisms of epilepsy and improve the HMM to generate data nearly identical to actual IEEG.

Overall the above analyses comprises the first application of evolutionary algorithms, frequent itemset mining, receiver operating characteristic curves, regression of probability density curves, Markov modeling, and multifactor analysis of variance in validation to establish a methodical solution for electrographically isolating, analyzing, and synthesizing appreciably epileptic brain.

5.2 Future Research

The following subsections briefly discuss the projected design of some future work stemming from the presented research.

5.2.1 Alternative Evolutionary Algorithms for Creating A-Features

Although this work strongly emphasizes the role of evolutionary algorithms to create and combine features, only two algorithms were considered among many other similar heuristics (e.g., grammatical evolution, shuffled frog leaping, ant colony optimization). A survey of these procedures applied to the data of this research to produce high-quality features would be interesting to compare results, trade-offs, and properties of each technique, considering the differences between each in stochastically traversing the fitness landscape.

5.2.2 Improving the HMM that Simulates Interictal IEEG

Improvements could be made to the presented HMM in order to generate data nearly identical to actual interictal IEEG by incorporating supplementary normal and abnormal electrographic waveforms (e.g., seizures, theta rhythms, spikes, sharp waves) and parameters to control the apparent SNR.

5.2.3 Dynamic Spatial Frequent Itemset Mining

This work presents the application of FIM to map putative epileptic networks in a static diagnostic mode, where an entire long-duration, multi-electrode record of interictal activity was processed to yield one cluster of considerably epileptogenic electrodes. Now

implement the same methodology in a dynamic diagnostic mode. That is, repeatedly run the FIM algorithm in small windows of time (several minutes) still over long durations and multiple electrodes with epileptiforms presumably bursting in a subset of sites, observing the results at each increment of executing the algorithm. Does this result converge (in a sense) to the result of applying the static diagnostic approach?

5.2.4 Frequent Itemsequence Mining

The work is only concerned with mining itemsets, but not the associations between items within an itemset. For instance, if a maximally frequent itemset for a particular subject equaled {2, 6, 7, 1, 4, 3}, with each number representing a channel in the cluster, perhaps the following was the sequence in which the cluster formed, which may hold some significance in understating the mechanisms of the epilepsy for the subject: {2, 7}, {2, 7, 4}, {2, 6, 7, 4} ⋯, {2, 6, 7, 1, 4, 3}. Thus, in this problem, mining itemsequences and their connections becomes important.

5.2.5 Correlation Analyses

Foremost, it is important to contrast an analysis of electrographic recordings in epileptic tissue using the presented methodology before resective surgery with clinical outcomes after treatment to investigate the following: 1) change the severity of seizures, if any transpire; 2) change in epileptogenic regions, if applicable; and 3) efficacy of therapy. This study would probe the existence of a relation between the chosen area for treatment and clustered electrodes in the analysis according to the abovementioned factors. The most favorable result of this study would demonstrate that resection of the

epileptic network as identified by the cluster made a patient seizure-free, proving the analysis key in completely pinpointing the epileptogenic zone.

Secondly, if the technical difficulties of simultaneously recording fMRI and EEG data with high quality can be resolved, then inspecting the moments in time when an epileptic oscillation changes the electrographic activity in the EEG and possibly the functional measures of the MRI would be an interesting examination to explain epileptic oscillations from a new viewpoint if any correlations exist between the two types of data.

Lastly, video recordings along with EEG may be helpful to observe if the patient experienced any auras or physical responses when an interictal epileptic event occurred, providing yet another standpoint on the epileptiforms. Perhaps a correlation between the magnitude of the response and a key parameter of the interictal event holds that can aid therapeutic means to control epilepsy.

Appendix A: List of Protected Functions

Table A-1 List of the protected functions, corresponding symbols and protection (if applicable) that was available to the genetic programming (GP) algorithm.

Function	Symbol	Protection
Addition	+	N/A
Subtraction	–	N/A
Multiplication	×	N/A
Division	÷	Output 0 when denominator input is zero
Square	$( )^2$	$( )^2$
Cube	$( )^3$	$( )^3$
Square root	$\sqrt{\lvert\ \rvert}$	Apply an absolute value operator before radical
Natural logarithm	ln	Output zero for an argument of zero; and Apply an absolute value operator to negative arguments
Absolute value	$\lvert\ \rvert$	N/A
Sine	sin	N/A
Cosine	cos	N/A
Arctangent	atan	N/A
Maximum	max	N/A
Minimum	min	N/A

Appendix B: Library of Classical Features

This appendix contains a list of conventional, or classical features (c-features) that were considered as input to the genetic programming (GP) procedure for the creation of artificial features that optimally discern specific epileptic oscillations from background activity in an EEG time-series. The measures in the ensuing library are based upon literature concerning seizure detection, previous work on detecting abnormal electrographic waveforms, and presumptions about the general physics of the waveforms to be detected. Though the below list is not an exhaustive list of quantities, new features can be added. The following notation is helpful to understand some of the listed features, which included four major analysis domains (i.e., time, frequency, statistics, information theory).

Let $\mathrm{n} = \{v : v \geq (i-1)\mathrm{N}+1 \wedge v \leq i\mathrm{N}, i \in \mathbb{N}\}$ be a set of time indices in samples and $\mathrm{f} = \{v : v \geq 0 \wedge v \leq \mathrm{Fs}/2\}$ be a set of frequency values in Hertz indices, where N is the number of elements in n, M is the number of elements in f, and Fs is the sampling rate for the EEG data. In addition, denote the IEEG sequence over a certain time interval *i* as *x*(n), where $\mathrm{x} = \{v : v_j = x(j) \forall j, 1 \leq j \leq \mathrm{N}\}$, the Fast Fourier Transform (FFT) of *x*(n) as X(f), and the normalized FFT of *x*(n) as $\widehat{\mathrm{X}}(\mathrm{f})$. Finally, let $x^{\mathrm{ord}}(\mathrm{n})$ define the ordered values of *x*(n), where $x^{ord} = \{x_1^{ord}, x_2^{ord}, \cdots, x_{\mathrm{N}}^{ord}\}$, and $pdf(x(\mathrm{n})) = pdf(\mathrm{x})$, where *pdf* is the probability density function (PDF) of a set of values.

Table B-1 Time Domain Features

Feature	Equation	Description
Energy	$\frac{1}{N}\sum_{n} x(n)^2$	Average instantaneous energy
Nonlinear Energy	$\frac{1}{N}\sum_{n}(x(n)^2 - x(n-1)x(n+1))$	Change in amplitude and frequency
Curve Length	$\frac{1}{N}\sum_{n}\lvert x(n) - x(n-1)\rvert$	Length of the curve, similar to arc length or curvature
Root Mean Square (RMS)	$\sqrt{\frac{1}{N}\left(\sum_{n} x(n)\right)^2}$	RMS amplitude
Mean Rectified Value	$\frac{1}{N}\sum_{n}\lvert x(n)\rvert$	Average magnitude
Crossings	$\frac{1}{N-1}\sum_{n\in\mathbb{N}\cap n>2}(\lvert(sign(x(n)-x(n_o))-sign(x(n-1)-x(n_o))\rvert>0)$	Average number of intersections across an amplitude reference

Table B-2 Frequency Domain Features

Feature	Equation	Description
Maximum PSD	$\max_{f}\lvert X(f)\rvert^2$	Maximum value of the estimated power spectral density
Mean PSD	$\frac{1}{M}\sum_{f}\lvert X(f)\rvert^2$	Mean value of the estimated power spectral density
Peak Frequency	$f_{max}\ni\lvert X(f_{max})\rvert=\max_{f}\lvert X(f)\rvert$	Frequency corresponding to the maximum value of the estimated power spectral density
Mean Frequency	$\frac{\sum_{f} f\cdot\lvert\hat{X}(f)\rvert^2}{\sum_{f}\lvert\hat{X}(f)\rvert^2}$	Frequency that is the centroid of power spectral density

Table B-3 Statistics Domain Features

Feature	Equation	Description
Mean	$\frac{1}{\mathrm{N}}\sum_{\mathrm{n}} x(\mathrm{n})$	Arithmetic mean or average of amplitude values
Standard Deviation	$\sqrt{\frac{1}{\mathrm{N}\text{-}1}\sum_{\mathrm{n}}(x(\mathrm{n})-\frac{1}{\mathrm{N}}\sum_{\mathrm{n}} x(\mathrm{n}))^2}$	Unbiased standard deviation of amplitude values
Skewness	$\frac{\frac{1}{N}\sum_{n}\left(x(n)-\frac{1}{N}\sum_{n} x(n)\right)^3}{\left(\sqrt{\frac{1}{N\text{-}1}\sum_{n}(x(n)-\frac{1}{N}\sum_{n} x(n))^2}\right)^3}$	Asymmetry about the mean amplitude value
Kurtosis	$\frac{\frac{1}{N}\sum_{n}\left(x(n)-\frac{1}{N}\sum_{n} x(n)\right)^4}{\left(\sqrt{\frac{1}{N\text{-}1}\sum_{n}(x(n)-\frac{1}{N}\sum_{n} x(n))^2}\right)^4}$	Affinity to outliers in distribution of amplitude values
Median	$\left\{\begin{array}{ll} x^{ord}(\frac{\mathrm{N}+1}{2}), & \mathrm{N\ is\ odd} \\ \frac{x^{ord}(\frac{\mathrm{N}}{2})+x^{ord}(\frac{\mathrm{N}}{2}+1)}{2}, & \mathrm{N\ is\ even} \end{array}\right\}$	Middle value when amplitude values are ordered from smallest to largest
Interquartile Range	$\left\{\begin{array}{l} median\{x^{ord}_{\frac{\mathrm{N}+1}{2}+1},\cdots,x^{ord}_{\mathrm{N}}\}-median\{x^{ord}_{1},\cdots,x^{ord}_{\frac{\mathrm{N}+1}{2}}\},\ \mathrm{N\ is\ odd} \\ median\{x^{ord}_{\frac{\mathrm{N}}{2}+1},\cdots,x^{ord}_{\mathrm{N}}\}-median\{x^{ord}_{1},\cdots,x^{ord}_{\frac{\mathrm{N}}{2}}\},\ \mathrm{N\ is\ even} \end{array}\right\}$	Spread of the amplitude values

Table B-4 Information Theory Domain Features

Feature	Equation	Description
Shannon Entropy	$-\sum_{x} pdf(x) \cdot \log_2(pdf(x))$	Level of disorder, uncertainty, or randomness
Spectral Entropy	$-\sum_{f} \lvert \hat{X}(f) \rvert \cdot \log_2(\lvert \hat{X}(f) \rvert)$	Level of disorder, uncertainty, or randomness
Reyi Entropy	$-\sum_{x} pdf(x)^p \cdot \log_2(pdf(x)^p)$	Level of disorder, uncertainty, or randomness
Complexity*	$\frac{\log_2(N)}{N} SequenceComplexity(x(n))$	Level of sequence complexity given by the number of unique subsequences within the sequence

*Note: The Complexity feature cannot be represented by a simple equation. In lieu of this difficulty, the algorithm *SequenceComplexity*(x(n)) was written but the code is omitted. For more details, see Zhang et al,"Detecting Ventricular Tachycardia and Fibrillation by Complexity Measure" IEEE 1999.

Appendix C: Genetic Programming Algorithm

C.1 Overview of the Algorithm

Genetic programming (GP), which was established by Koza [147-149], is a general-purpose, global search optimization procedure that computes an optimal solution—in the sense of an objective function—by imitating four processes of natural evolution described by Darwin: *selection*, *crossover*, *mutation*, and *survival*. More specifically, the algorithm initializes a set (*population*) of solutions (*individuals*) of size *P* with each element representing a mathematical operation on the input of the GP in the form of a tree structure; computes per individual a measure of *fitness* to an ideal solution, or a value for the objective function; and executes the genetic operations (i.e., selection, crossover, mutation, survival) to create new populations while either maximizing or minimizing the fitness of the best *individual* (final solution), depending on the formulation of the problem.

Although many variations of the genetic operations in GP exist, the basic concept of each process is constant. In the selection stage, a subset of the current population (*intermediate population*) is chosen based upon the fitness of the individual; in the crossover stage, the GP creates new individuals (*children*) using combinations of a pair of individuals (*parents*) from the intermediate population that will belong to a new population; the mutation stage introduces diversity into the new population by randomly altering the makeup of a subset of individuals in the new population; and the survival stage simply selects the fittest individuals from the new population, creating a new initial

population of size P for subsequent iterations of the GP. The GP algorithm terminates upon reaching a predefined number of iterations (*generations*) or a level of fitness among the present population. The following figures elucidate the operation and application of a GP. Figure 3.1-1 depicts the overall genetic programming procedure; Figure 3.1-2 is an example of a result from a GP program for a particular problem; and Figures 3.1-3 and 3.1-4 demonstrate the crossover and mutation operations, respectively.

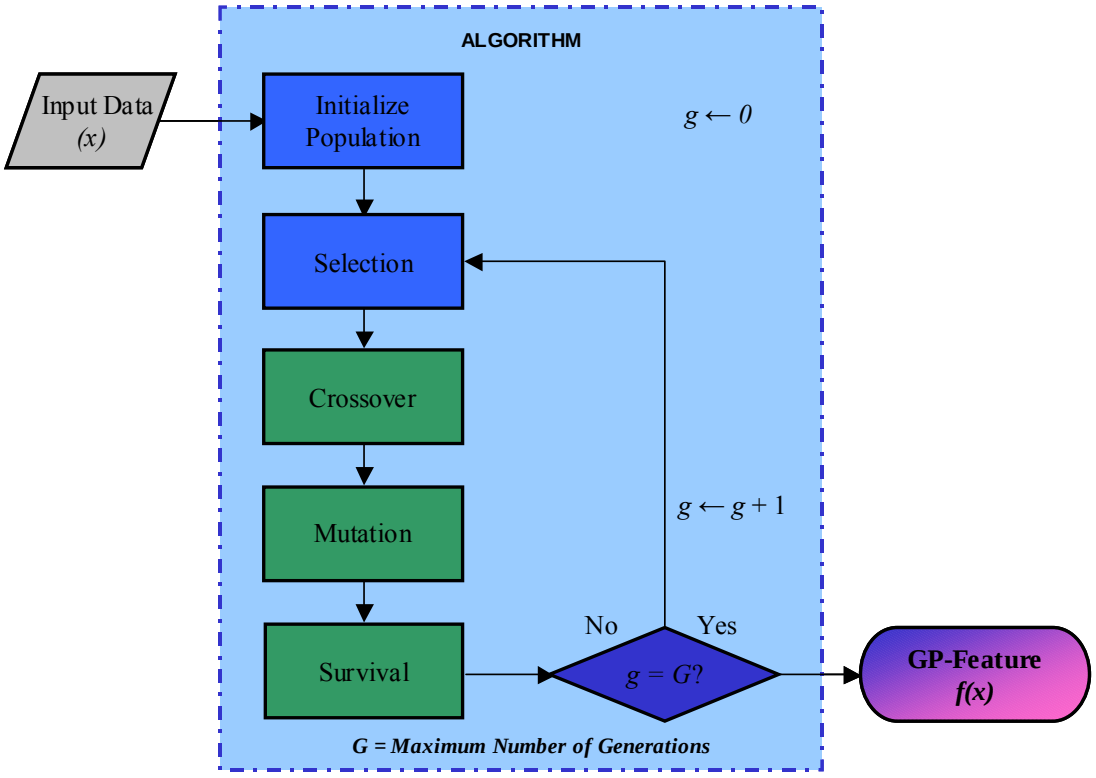


Figure C.1-1 This figure illustrates the steps of a canonical genetic programming (GP) algorithm. The variable g denotes the number of generations, or iterations of the procedure, while the constant G is the desired number of generations. The result of the GP is a function—typically nonlinear—of the input data. The fitness measure depends upon the GP application.

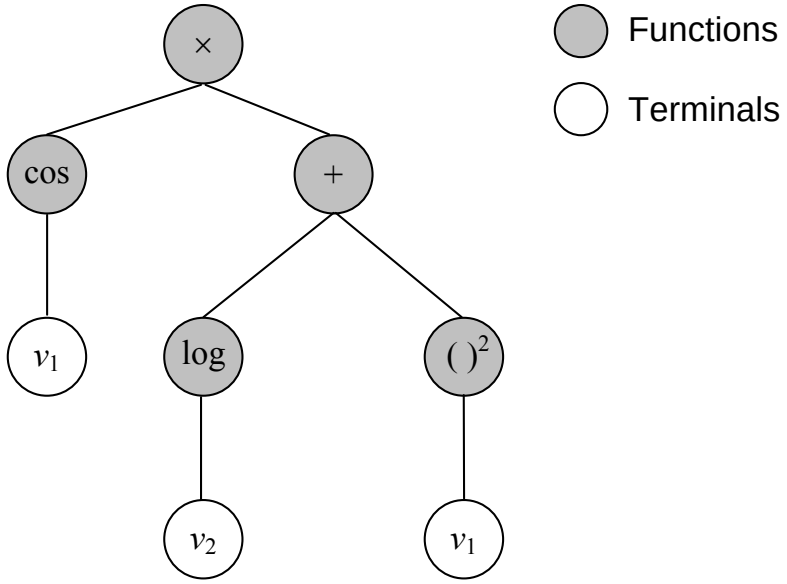


Figure C.1-2 This figure exemplifies a solution from a particular GP procedure. The tree structure represents the following expression: $\cos(v_1) \cdot (\log(v_2) + (v_1^2))$.

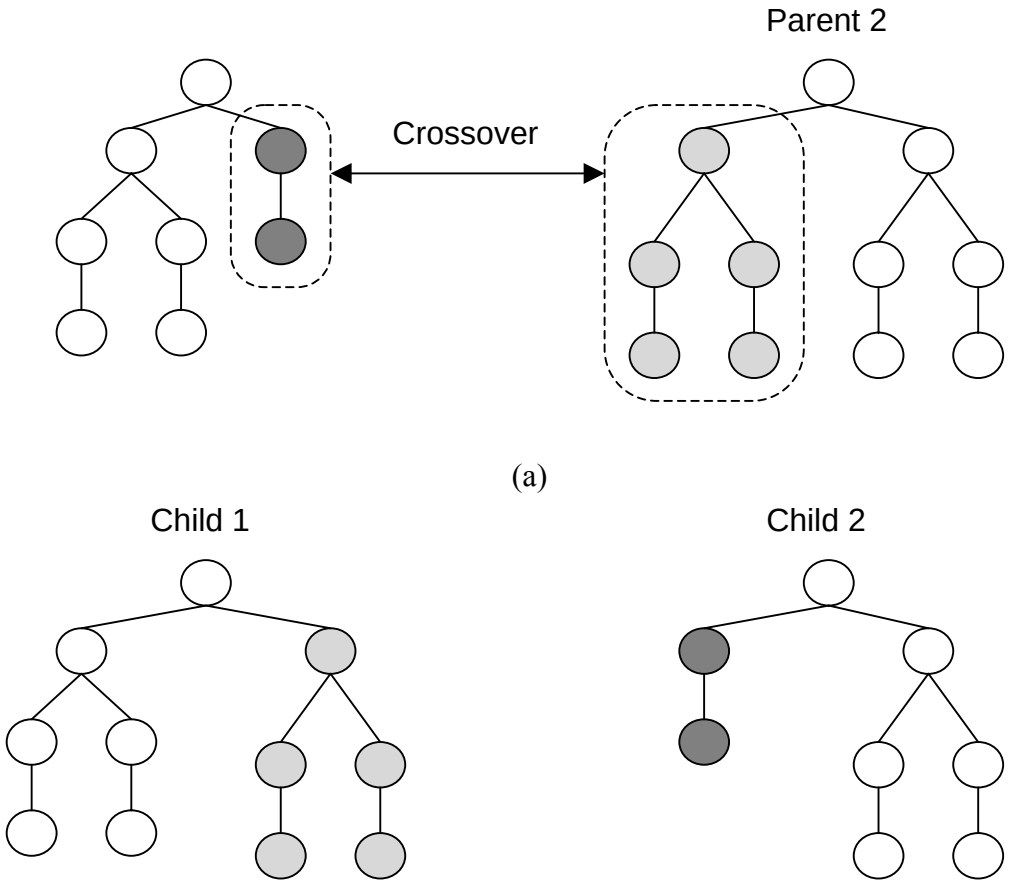


Figure C.1-3 This figure demonstrates a typical crossover operator in a GP procedure: a) the parents swap sub-trees at randomly selected nodes from each parent tree; and b) the children are the result after crossover is performed.

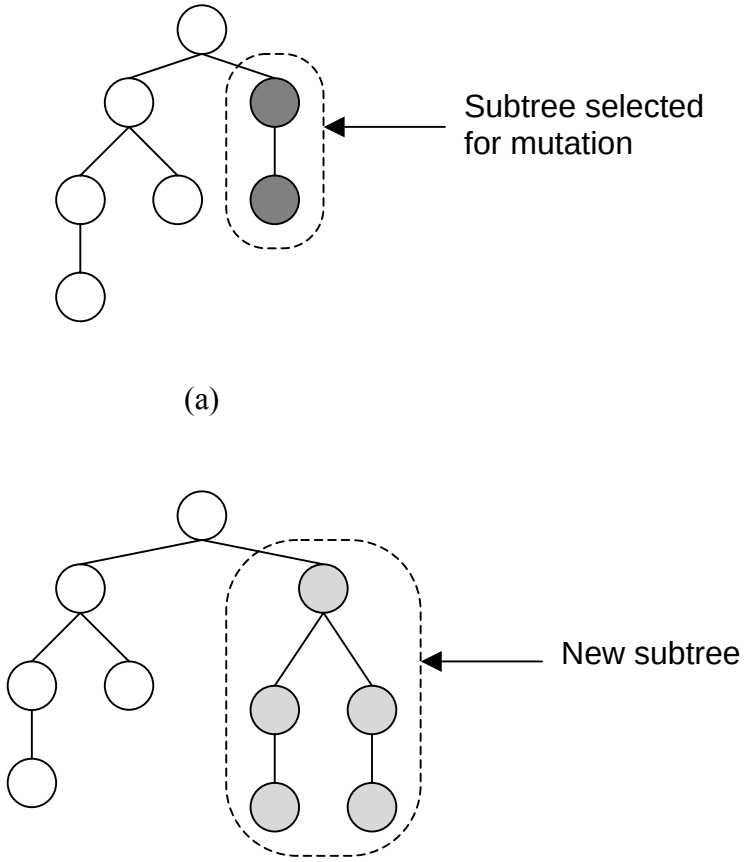


Figure C.1-4 This figure demonstrates a typical mutation operator in a GP procedure: a) a sub tree is randomly selected from the node of a parent tree; and b) altering the sub tree of the parent produces a child.

Each individual in any population is a tree structure, similar to the output of the GP illustrated in Figure 1.3.1-2, that contains two types of nodes: *functions* that receive a number of arguments on which they operate and *terminals*, which are values selected from the input. The functions are simple mathematical operations (e.g., addition, multiplication, logarithm, sine) that have been protected to avoid returning undefined values. In the framework of feature creation, the terminals are some transformation of an original or enhanced time-series such as an n-point delay or computation of classical or physics-based features (e.g., mean energy, mean entropy, kurtosis).

C.2 Application to Creating Artificial Features

A genetic program can be applied for the creation of an artificial feature that discriminates between epileptic events and background activity (non-epileptic events) within IEEG. Consider the two-class problem of this work. The input to the GP is a matrix, **X**, containing either delays in the IEEG time-series in which each column is a delayed time-series or feature-series in which each column is a series from an extracted c-feature; and a column-vector of labels, **c**, with the same number of rows as **X** in which each label from the set {0, 1} signifies the class—one for an epileptic event and zero otherwise—to which a corresponding row-vector of **X** belongs.

When **X** is a matrix of delayed time-series, or essentially a convolution matrix, the implementation is referred to as GP-DD. In GP-DD, $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_{N\text{-}1}, \mathbf{x}_N]$, where N is the number of embedded delays in the original time-series, $x[n] = \mathbf{x}_0$, and $\mathbf{x}_i = x[n-i],\ 0 \le i \le N$ describes the i^{th} column-vector of **X**. In GP-CF, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_{F\text{-}1}, \mathbf{x}_F]$, where F is the number of c-features to extract from in the original

time-series, $x[n]$, and $\mathbf{x}_j = \Phi_j\{x[n]\}$, $0 \le i \le F$ describes the j^{th} column-vector of **X** with each feature, Φ_j, sharing the same values for the sliding window, *W*, and shift in the sliding window, *D*.

In either case, the GP finds an optimal tree, or a-feature, that consists of two components: 1) a multiple-input, single-output transformation, *f*, using the functions in Appendix A; and 2) terminals equal to column-vectors selected from **X**. In this application, where Equation C.2-2 represents the objective function, optimality means that the GP returns a transformation and set of terminals to maximally distinguish between epileptic and non-epileptic events. Equation C.2-1 conveys the input-output transformation that results in creating an artificial feature; where **y** is a column vector with the same number of rows as **X** and the values of the artificial feature, and $\hat{\mathbf{X}}$ is a matrix with the selected columns from **X**.

$$\mathbf{y} = f(\hat{\mathbf{X}}) \tag{C.2-1}$$

$$fitness = (1 - \int \min_{y}(p(\mathbf{y} \mid \mathbf{c} = 0), p(\mathbf{y} \mid \mathbf{c} = 1))\, dy) \cdot \frac{\left|\mu_1 - \mu_0\right|}{\sqrt{(\sigma_1^2 + \sigma_0^2)/2}} \tag{C.2-2}$$

In Equation C.2-2, $p(\mathbf{y} \mid \mathbf{c} = k)$ is the prior distribution of an a-feature in class $k \in \{0,1\}$ with mean μ_k and standard deviation σ_k.

Appendix D: Particle Swarm Optimization Algorithm

D.1 Overview of the Algorithm

Akin to genetic programming and genetic algorithms, particle swarm optimization (PSO) is a global search algorithm on the family-tree of evolutionary algorithms that computes an optimal solution according to a fitness (objective) function by emulating a naturally occurring process. Devised by Kennedy and Eberhart [161, 162] in the mid-1990's, PSO is inspired by the social behavior of a flock of birds searching a landscape for the best location to eat: each member of the flock communicating good locations to the other birds while quickly adapting their position and speed to reach the best location according a leader. For the PSO algorithm, a *swarm* (flock) of *particles* (birds) traverses a *fitness landscape* (landscape) for an optimal solution (location) according to the particle with the best fitness of the entire swarm and a few particles, each with the highest fitness among a subset of the swarm.

Let the vectors $\mathbf{x}_i \in \mathbb{R}^m$ and $\mathbf{v}_i \in \mathbb{R}^m$ respectively represent the position and velocity, or rate of position change, of the i^{th} particle in a swarm of n particles. Furthermore, let the vectors $\mathbf{y}_i \in \mathbb{R}^m$ and $\mathbf{g} \in \mathbb{R}^m$ respectively represent the best position by an individual (i^{th}) particle (local best fitness) and the best position associated with the leader of the swarm (global best fitness). The following formulae express how each particle updates its position and velocity after each iteration of the algorithm, k.

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) + \mathbf{v}_i(k) \tag{D.1-1}$$

$$\mathbf{v}_i(k+1) = w \cdot \mathbf{v}_i(k) + c_1\mathbf{r}_1(k) \bullet (\mathbf{y}_i(k) - \mathbf{x}_i(k)) + c_2\mathbf{r}_2(k) \bullet (\mathbf{g}(k) - \mathbf{x}_i(k)) \tag{D.1-2}$$

Equation D.1-1 is referred to as the *position update equation* and Equation D.1-2 is known as the *velocity update equation*. On the k^{th} iteration of PSO, after a new position and velocity are calculated for each particle, the fitness due to the new position is compared against the global best fitness and local best fitness of the $(k\text{-}1)^{\text{th}}$ iteration. If there is an improvement on the global (local) fitness at the new position, then $\mathbf{y}_i$ ($\mathbf{g}$) is also updated. The PSO procedure continues until a stop condition is met, denoting convergence to an optimal solution.

Commonly, the PSO algorithm is initialized with the following conditions: $\mathbf{x}_{ij} \in U[a_j, b_j]$ and $\mathbf{v}_i = \mathbf{0}$, $\forall i, 1 \le i \le n$ and $1 \le j \le m$, where a_j and b_j are the limits of the search space in the j^{th} dimension [180]. Moreover, in the k^{th} iteration of the PSO, $\mathbf{v}_{ij}(k)$ is clamped [180] to the range [$-V_{max}$, V_{max}] to keep the values of $\mathbf{x}_{ij}(k)$ within the range $[a_j, b_j]$. The inertia weight, w, is a factor used to control the balance of the search algorithm between global and local exploration of the swarm due to the impact of the previous velocities on the current velocities [13, 165, 176]. The coefficients c_1 and c_2 are the *acceleration coefficients*, where the first coefficient moderates the maximum step size of the swarm towards a good local position and the second coefficient moderates the step size of the swarm toward a good global position. Both coefficients are bounded by the inequality $0 < c_1, c_2 < 2$. The random variables $r_1(t)$ and $r_2(t)$, each of which is uniformly distributed over the range from zero to one, make the PSO algorithm stochastic so that

each particle possesses a level of randomness in its movement (representing the explorative capability of a single bird in the flock) while the swarm shares the influences of the individual.

D.2 Application to Creating Artificial Features

Unlike the genetic programming algorithm, which manipulates a tree-like structure with heuristics, the particle swarmed optimization algorithm lacks a definite structure on which to operate its heuristics. A very complementary structure with which to combine the PSO is a neural network (NN), which can effectively approximate an input-output relationship by iteratively adjusting the connections (weights) between elements (summations and activation functions) within layers of the network. In fact, the *universal approximation theorem for neural networks* states that every continuous function that maps intervals of real numbers from some input to an output can be closely approximated by a multiple layer NN with only a single hidden layer for certain classes of activation functions (e.g., sigmoidal) [175, 182]. Consequently, this work uses two hidden layers of sigmoidal activation functions. The composite process of the NN and the PSO is referred to as PSO-NN and is illustrated in Figure D.2-1.

The input to the PSO-NN is a matrix, **X**, containing embedded delays in the IEEG time-series, $x[n]$, in which each column is a delayed time-series. Essentially, **X** is a convolution matrix, $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_{N\text{-}1}, \mathbf{x}_N]$, where N is the number of embedded delays in the original time-series, $x[n] = \mathbf{x}_0$, and $\mathbf{x}_i = x[n-i]$, $0 \le i \le N$ describes the i^{th} column-vector of **X**. Recalling the terminology of PSO in the previous section, each weight is the position, and the change in the value of a weight is the velocity; the

collection of weights is the particle, and the set of particles is the swarm. The PSO finds an optimal solution, or an a-feature, that consists of two components: 1) the matrix of weights for the neural network, ϕ; and 2) selected column-vectors from **X**. In this application, where Equation D.2-2 represents the objective function, optimality means that the PSO returns a neural network and set of delays in $x[n]$ to maximally distinguish between epileptic and non-epileptic events. Equation D.2-1 conveys the input-output transformation that results in creating the j^{th} value in the feature-series of an artificial feature, **y**.

$$\mathbf{y}_j = \phi\left(\mathbf{x}_j\right) = g\left(\mathbf{w}_L g\left(\mathbf{w}_{L-1} \cdots g\left(\mathbf{w}_l \cdots g\left(\mathbf{w}_1 \mathbf{x}_j\right)\right)\right)\right) \qquad \text{(D.2-1)}$$

In Equation D.2-1, $\phi: \mathbb{R}^m \to \mathbb{R}$ is a function representing the neural network, $\mathbf{w}_l$ is the matrix of weights for the l^{th} layer of the NN, g is an activation function, and the vector $\mathbf{x}_j$ represents the j^{th} row of **X** along with a bias term of one for the NN, or $\mathbf{x}_j = [1, x[(j-1)-0], x[(j-1)-1], \cdots x[(j-1)-(N-1)], x[(j-1)-N]]$. The artificial feature, **y**, is a column vector with the same number of rows as **X**. The PSO is driven by the objective function, or fitness, described in Equation D.2-2.

$$fitness = (1 - \int \min_y (p(\mathbf{y} \mid k=0), p(\mathbf{y} \mid k=1))\, dy) \cdot \frac{\left|\mu_1 - \mu_0\right|}{\sqrt{(\sigma_1^2 + \sigma_0^2)/2}} \qquad \text{(D.2-2)}$$

In Equation D.2-2, $p(\mathbf{y}\,|\,k)$ is the prior distribution of an a-feature, $\mathbf{y}=\phi(\mathbf{x})$, in class $k \in \{0,1\}$ with mean μ_k and standard deviation σ_k.

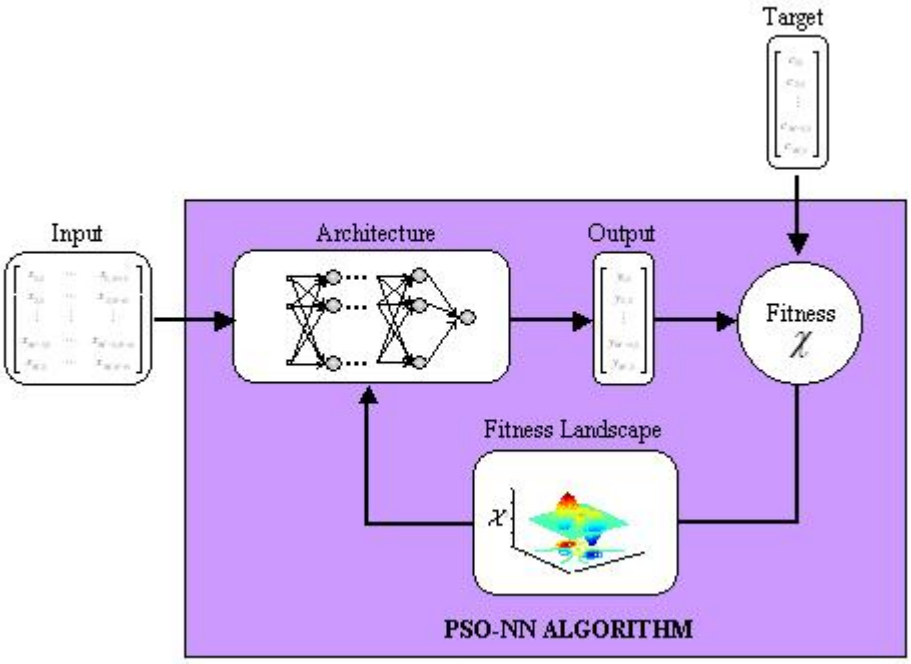


Figure D.2-1 This figure illustrates the training of a neural network via supervised learning. Given an *a priori* input and corresponding target, a learning algorithm adjusts the weights of the NN until a stopping condition is met after comparing the output of the NN and the target.

Appendix E: Frequent Itemset Mining Algorithm

E.1 Overview of the Algorithm

Although there are many applications of data-mining, frequent itemset mining (FIM) is the problem most relevant to this work. Agrawal and colleagues first present the topic of FIM about a decade ago, developing perhaps the most established procedure for FIM, APRIORI [155, 158]. Since the seminal efforts of Agrawal et al., much work has been done to improve the efficiency and reduce the consumption of memory for FIM algorithms. However, the most noteworthy research that has advanced the implementation of APRIORI belongs to Ferenc Bodon, who describes an implementation of APRIORI that outperforms other implementations, including the original algorithm [153, 155].

Bodon's scheme is based upon a data structure known as a *trie*, which was originally introduced by de la Briandais and Fredkin [156], to efficiently store and retrieve words of a dictionary and is essentially similar to a *tree* structure. That is, the trie has depth such that a node at the top of the tree (depth 0) is the root and nodes at other depths are connected by edges (links). If a node at a higher depth points to a node at a lower depth, the higher node is referenced as the *parent* of the lower node and the lower node is referenced as the *child* of the higher node. A child at the lowest depth in the trie is a *leaf*. Before the usefulness of this structure for FIM is explained, a few terms will be discussed. These terms are consistent with those introduced by Agrawal et al.

E.2 Definitions for the Framework of FIM

1) An *item* a_k refers to an integer k, $k \in [1, N]$, corresponding to some physical factor (variable) of the problem to which an FIM is applied.
2) An *n-itemset* is a set of n items.
3) An *n-itemsequence* is an ordered set of items (or an ordered itemset).
4) An m-itemset (m-itemsequence) $\{b_1, \ldots b_m\}$ is *contained* in another n-itemset (n-itemsequence) $\{a_1, \ldots a_n\}$ if $\exists\, i_j \in [1, \ldots n],\ j \in [1, \ldots m],\ m \le n,\ \ni\ b_j = a_{i_j}\ \forall\, j$.
5) A *database* is a list of n-itemsets (n-itemsequences) with possibly different values for n per itemset (itemsequence) within the list.
6) The *count* of an itemset (itemsequence), which is noted as $count_\Sigma(I)$, is the number of itemsets (itemsequences) in a database Σ containing the itemset (itemsequence) I.
7) The *support* of an itemset (itemsequence), which is noted as $support_\Sigma(I)$, is the fraction of itemsets (itemsequences) in a database Σ containing the itemset (itemsequence) I. Thus the support can be defined via the following expression: $support_\Sigma(I) = \dfrac{count_\Sigma(I)}{|\Sigma|}$, where $|\Sigma|$ is the length of the database (or equivalently, the total number of itemsets in the database) and $0 \le support_\Sigma(I) \le 1$.
8) A *λ-frequent* itemset or itemsequence I possesses a support such that $\lambda \le support_\Sigma(I),\ 0 \le \lambda \le 1$. When I is a *maximal λ-frequent* n-itemset

(*n*-itemsequence), all (*n*+1)-itemsets (itemsequences) containing it are not *λ-frequent*.

Returning to the role of the trie structure in the new APRIORI implementation by Bodon, each node of the trie is a λ-frequent itemset, which represents a path from the root to the corresponding leaf. In addition to the itemset, a node stores the length *n* and support of the itemset. The links associating parent and child store the λ-frequent item in the itemset of the new child that is added to the itemset of the child's parent. Hence, the trie grows as APRIORI iteratively returns a new *n*-itemset that is built from a λ-frequent itemset, and the new itemset is retained to update the trie node if the itemset is λ-frequent. The procedure continues until a maximal λ-frequent n-itemset is discovered.

E.3 Application to Mining Epileptic Oscillations

Frequent itemset mining scheme is apt for investigating whether interictal/preictal epileptic oscillations throughout many IEEG electrodes can identify key electrodes that represent considerably dysfunctional parts of an epileptic brain and, in turn, map a network of epileptic activity that may be involved in ictogenesis. In the framework of FIM, given a record with multiple channels (electrodes) over a long period of time in which many interictal/preictal epileptic oscillations occur, an item is merely a number representing an electrode site in which the epileptic events during a predetermined interval of time have been detected and quantified (e.g., the total energy, average frequency), and an *n*-itemset is then a set of *n* channels. The database, Σ, is simply a temporal log of sets of channels. The length of the database, $T = |\Sigma|$, is the number of

intervals of time with corresponding *n*-itemsets in the record. Finally, a maximal λ-frequent *n-itemset* at the output of the algorithm for FIM signifies a group of channels—each channel exhibiting a high concentration of events—that occurs frequently with a likelihood of λ and merits significant consideration in mapping epileptic networks.

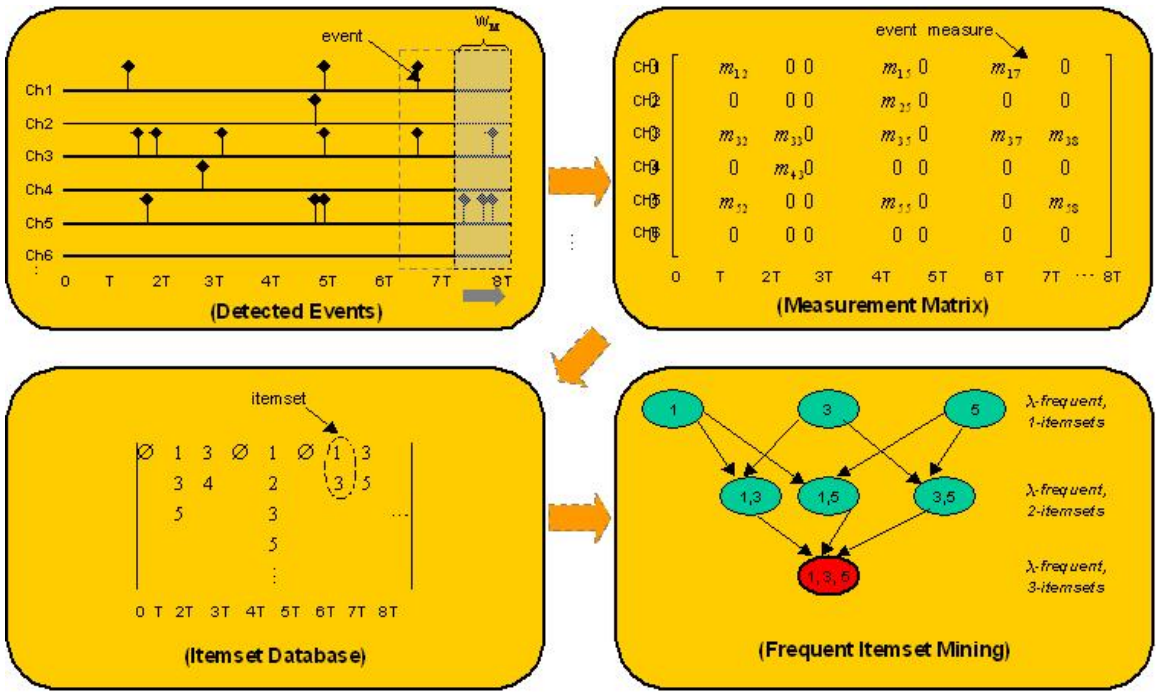


Figure E.3-1 This figure illustrates the application of frequent itemset mining for localizing concentrations of epileptic oscillation that may map an epileptic network.

Appendix F: Additional Experimental Results

This appendix contains the results or comments regarding a few experiments that were performed to support the application of evolutionary algorithms for feature extraction features, the chosen parameters for classification, and alternative features for the spatial-temporal matrix of measures in frequent itemset mining.

F.1 Evolutionary Algorithms for Extracting Features

This section briefly provides answers to the following questions concerning the application of an evolutionary algorithm (EA), in particular a genetic program (GP) because of the success presented in Section 4.2, for yielding subject-specific features in the presented detector of epileptic oscillations:

- How does the GP compare with a few other techniques for automatically selecting features relative to performance (i.e., sensitivity, specificity)?
- What trade-offs exist in choosing a classifier-independent or classifier-dependent fitness for the GP?

F.1.1 Comparing Techniques for Selecting Features

This experiment compares the performance of selecting and combining features via a genetic program, referred to as GP-CF, against four other approaches. The literature for detecting epileptic oscillations does not present a benchmark for automatically selecting features—as visual inspection is the standard, however the following techniques are considered adequate candidates for comparison:

- M1 – sequential forward selection (SFS). Sequential forward selection is a greedy search algorithm that determines an optimal subset of features for extraction by starting from an empty and sequentially adding a single feature in the superset to the subset if it improves an objective function.
- M2 – sequential backward selection (SBS). Sequential backward selection is similar to SFS. Instead, the search initializes with the full superset of features and removes a single feature that improves (or minimally worsens) an objective function to obtain the final optimal subset of features.
- M3 – principal component analysis (PCA). Typically, PCA is used to project multidimensional data in a smaller dimensional space by constructing an orthogonal basis with principal components. Making an assumption that the components hold some information useful for discriminating the data in different classes—which is not guaranteed [166], the coefficients of the principal component with the most information for representing the data determines the ranking of the features.
- M4 – genetic programming (GP). After a GP is executed to create an a-feature, only the terminals (selected features) are considered, ignoring the corresponding tree (nonlinear transformation) for the features.
- M5 – genetic programming of conventional features (GP-CF). The corresponding transformation of the features selected in M4 is performed.

The above five techniques are statistically tested for similarities in performance (i.e. sensitivity, specificity) across six subjects using an ANOVA-2 and a Tukey's test for multiple comparisons. In the ANOVA-2, the subject serves as the first factor (5 degrees of freedom) and the method (3 degrees of freedom) serves as the second factor, a total of 23 degrees of freedom (or the number of subjects times the number of methods minus one). The null hypothesis for each factor is that all methods to detect epileptic oscillations performed equally in terms of sensitivity and specificity. Considering each the sensitivity and the specificity as test-statistics, the ANOVA-2 returns P-values equal to zero, substantiating evidence that the null hypothesis is false. The Tukey's test, as illustrated in Figure F.1-1 and Figure F.1-2, reveal that the sensitivity of the methods differed statistically in a few cases (i.e., E02, E05, E09) and the specificity of GP-CF proves considerably better across patients.

Because of the great disparity in specificity between approaches and the moderate differences in sensitivity, selecting features using a genetic program is the final chosen technique. However, the results in the below figures seem to suggest that a *No Free Lunch Theorem* [166] across subjects. That is, that one scheme to select features given the data in one patient does not guarantee success in the set of data for another patient.

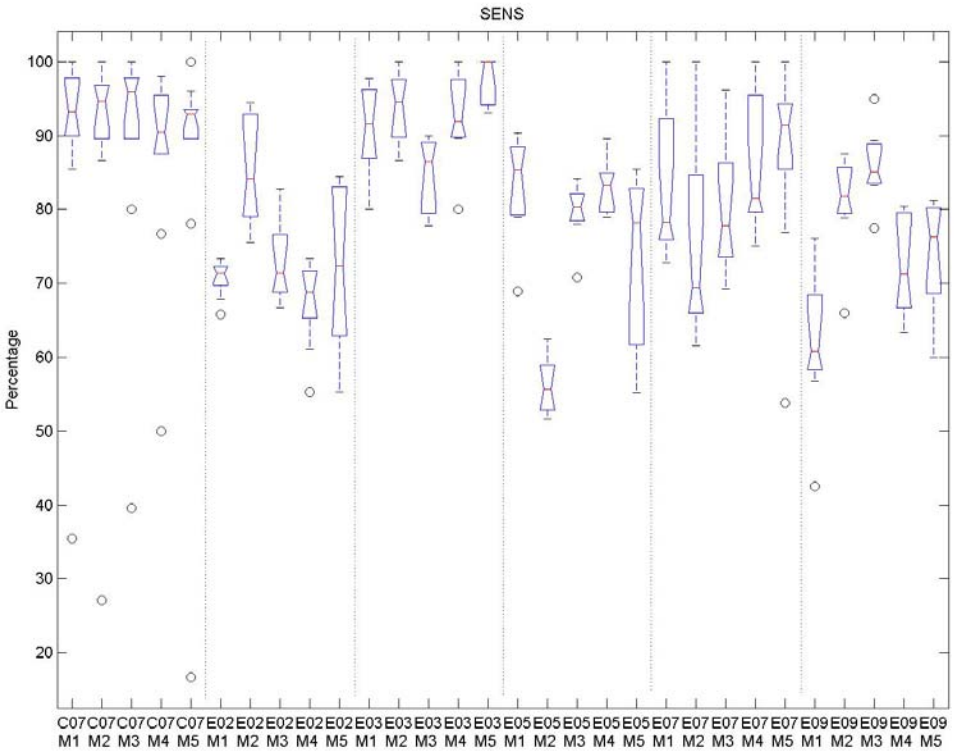


Figure F.1-1 This figure shows that the approaches to select features differed statistically in a few subjects (i.e., E02, E05, and E09), where either SFS (M1) or SBS (M2) exhibited better sensitivity than GP-CF (M5).

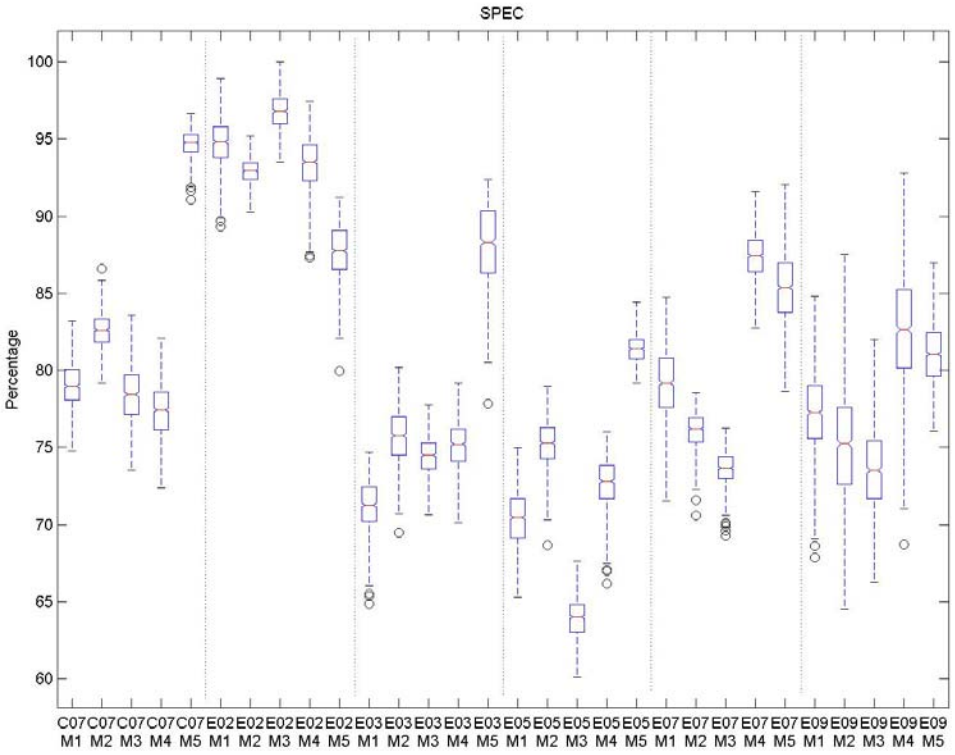


Figure F.1-2 This figure shows that the approaches to select features differed statistically across all subjects in terms of specificity. In fact, in five of the six cases, GP-CF (M5) or GP (M4) exhibited considerably better specificity than the alternatives M1-M3.

F.1.2 Choosing a Fitness Function for Creating Features

Evolutionary algorithms are implemented to create subject-specific artificial features that highly discriminate two classes of events. For this application, the EA must rely on a fitness function, or an objective function, that measures the ability to discern the binary classes. A traditional choice for a fitness function in this problem is the accuracy (or error) of classification using the feature as input to a classifier. However, this approach can be computationally expensive, considering that the EA searches a population (solution-space) of hundreds to thousands of individuals (potential solutions) for several iterations. As an alternative, the classifier-independent fitness defined by Equation F.1-1 is used. It is recognized that with such an objective function, the EA will search for a solution that slightly differs from the manner in which the a-feature is used, namely as input to a specific classifier, but the EA could also find a solution that is equivalent (or a close approximation) to one if a classifier-based fitness function is used. This notion is illustrated in Figure F.1-3, in which the receiver operating characteristic (ROC) curves of a system using three a-features: two are computed with a classifier-independent objective function and one is computed with a classifier-based objective function.

$$fitness = (1 - \int \min_{y}(p(\mathbf{y} \mid k=0), p(\mathbf{y} \mid k=1))\, dy) \cdot \frac{\left|\mu_1 - \mu_0\right|}{\sqrt{(\sigma_1^2 + \sigma_0^2)/2}} \qquad \text{(F.1-1)}$$

In the above expression, $p(\mathbf{y} \mid k)$ is the prior distribution of the created a-feature in class $k \in \{0,1\}$ with mean μ_k and standard deviation σ_k.

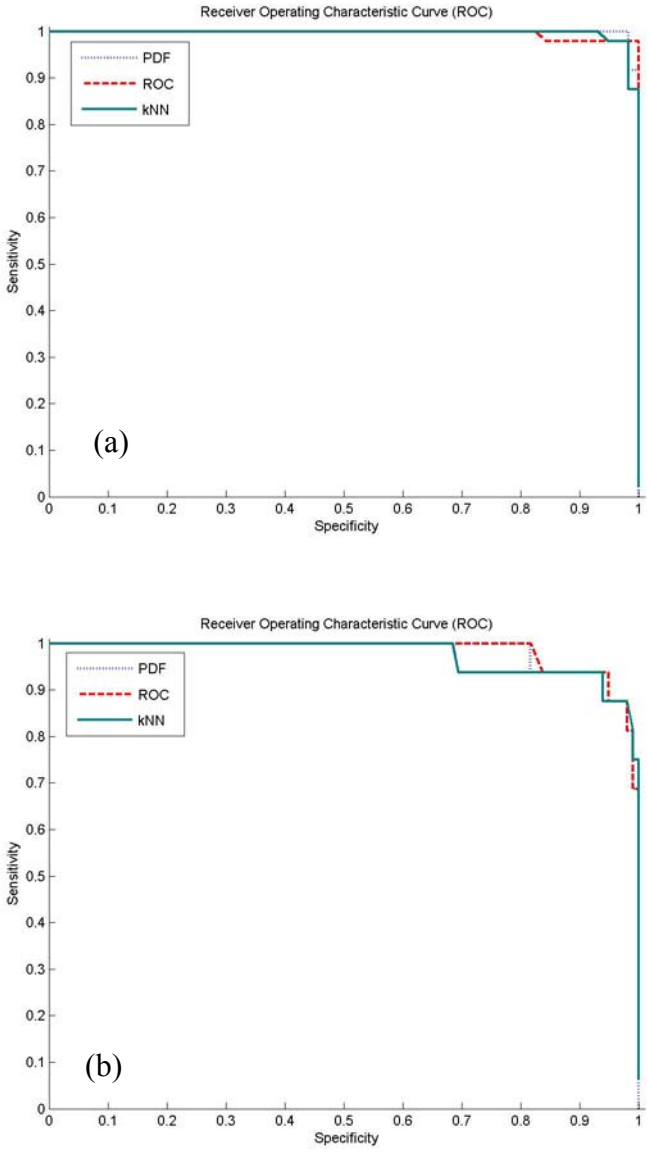


Figure F.1-3 This figure shows the ROC curves for artificial features that are created with a GP using classifier-independent (ROC, PDF) and classifier-dependent fitness functions for two subjects. In both cases, the "classifier-independent features" appear to be good approximations to the "classifier-dependent features."

F.2 Classification via *k*-Nearest Neighbor Rule

This section briefly provides answers to the following questions concerning the choice of a *k*-nearest neighbor (*k*-NN) rule for the classifier in the presented detector of epileptic oscillations:

- How does the *k*-NN rule compare with benchmarks relative to performance (i.e., sensitivity, specificity) and computational burden?
- How is the number of neighbors, *k*, determined and what value for each subject?
- How does the implementation of the *k*-NN classifier account for potential mismatches between the training and testing samples in terms of the proportion of points stratified by class in each sample? For instance, if the *k*-NN is "trained" with a sample possessing 50% of the points from the class for the epileptic oscillations and the remaining half of the sample for the complementary class but is tested on a sample possessing 80% (20%) of the points from the class for the epileptic oscillations (complementary class), what adjustments—if any—to the classifier is made?

F.2.1 Comparing the k-Nearest Neighbor Rule with Benchmarks

This experiment compares the performance and computational burden of five approaches to classifying epileptic oscillations (Class 1) and background activity (Class 0) using an artificial feature and chooses one that is apt for the detector: a *k*-NN, a fuzzy c-means (FCM) classifier, a probabilistic neural network (PNN), a support vector machine (SVM), and a threshold (tau). Literature demonstrates that the SVM and PNN are popular and successful architectures for classifying biomedical signals [62, 74]. An

FCM classifier represents the classical unsupervised techniques for classification, while the threshold represents the "standard" for detecting epileptic oscillations.

An analysis of variance (ANOVA) and Tukey's test compare the performance of the classifiers. First, a single feature-vector is extracted from an IEEG signal with values corresponding to Class 0 and Class 1. Second, a total of 15 Monte Carlo simulations of tuning and testing the classifiers is executed in which each simulation randomly draws two distinct random samples from the feature-vector—a training sample of 318 values and a testing sample of 1978 values—and records the sensitivity and specificity of each classifier. Third, the fifteen measures of sensitivity and specificity serve as statistical samples for the ANOVA. Figure F.2-1 and Figure F.2-2 illustrate the results of the ANOVA. This analysis demonstrates that the *k*-NN and threshold performs as well as the PNN and SVM in terms of specificity but occasionally greater than the two alternatives in terms of sensitivity.

To evaluate the computational burden, four trials of 15 Monte Carlo simulations of tuning and testing the classifiers is executed in which each simulation randomly generates two samples and each trial increases the sizes of the samples and registers the mean time of training and testing. Figure F.2-3a and Figure F.2-3b capture the average time needed to tune and test each classifier, respectively. Excluding the results of FCM because of poor performance, this analysis demonstrates that the *k*-NN, which merely stores the training samples, possesses the least burden for tuning and practically equals the PNN—though worse than the threshold—in the computational cost for testing.

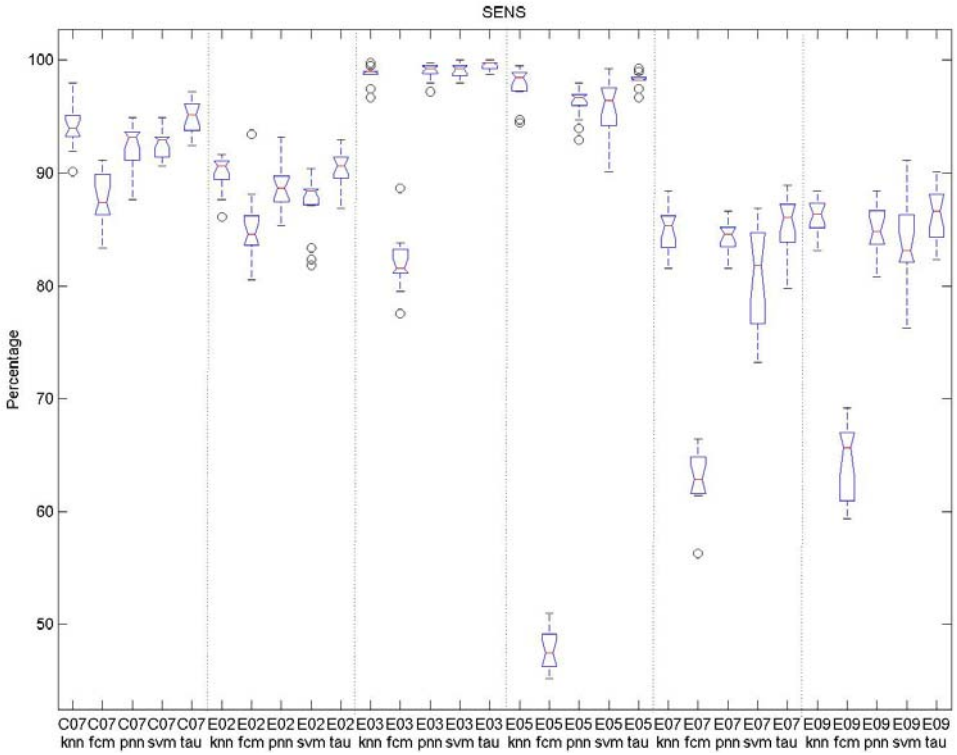


Figure F.2-1 This figure shows that in terms of the metric sensitivity across subjects, the *k*-NN and the threshold performed as well as and occasionally better than the SVM or PNN, while the FCM performed poorly.

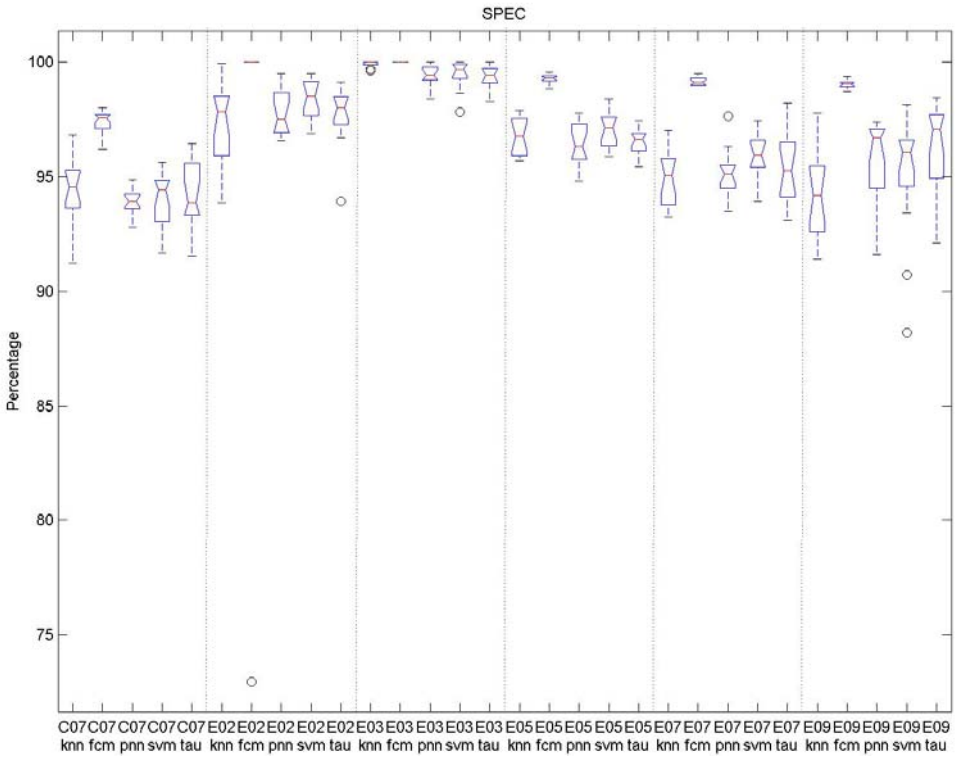


Figure F.2-2 This figure shows that in terms of the metric specificity across subjects, the *k*-NN and the threshold performed as well as the SVM or PNN, E09 notwithstanding for the *k*-NN, while the FCM typically performed better than the other classifiers.

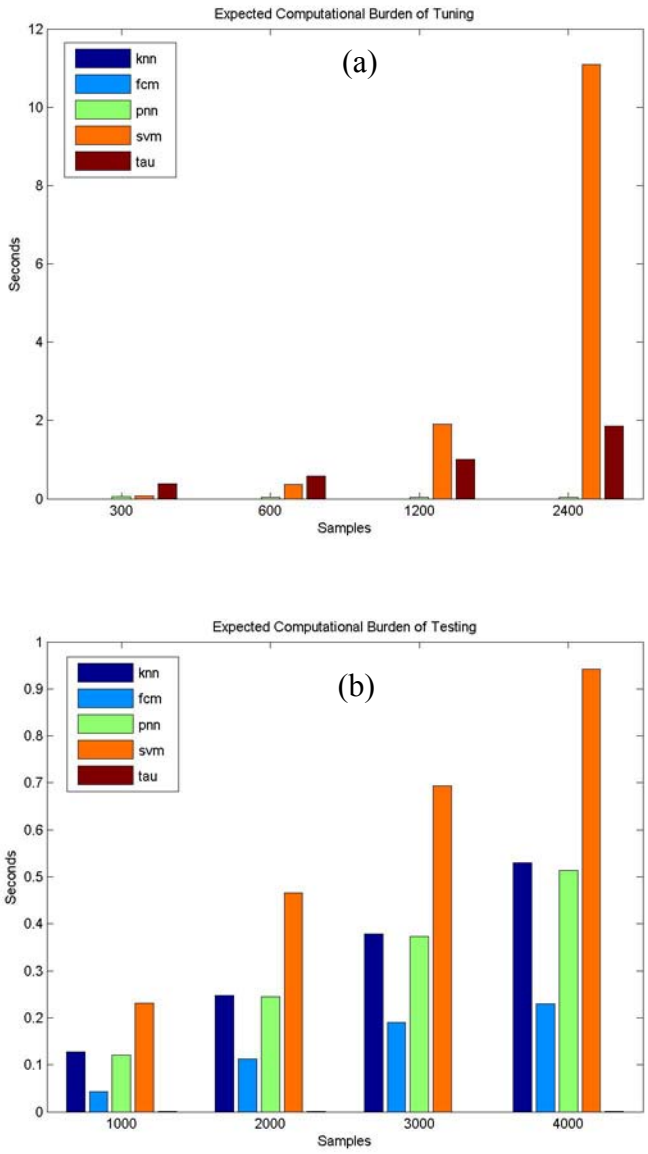


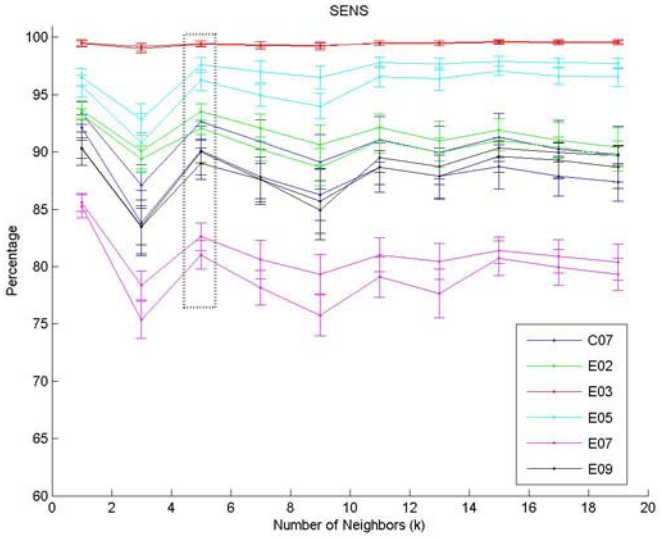
Figure F.2-3 This figure shows that the computational burden of training (*a*) and testing (*b*) four types of classifiers (*bars from left to right*): *k*-NN, FCM, PNN, SVM, and TAU. The unsupervised technique, FCM, and the *k*-NN required no time to train and relatively low time to classify. The expected computational burden of the PNN nearly equaled that of the *k*-NN, which possessed the edge in training time. The threshold essentially took no time to classify a feature-series, despite its duration, while the SVM took the longest amount of time to evaluate in training and testing.

F.2.2 Setting k in the k-Nearest Neighbor Rule

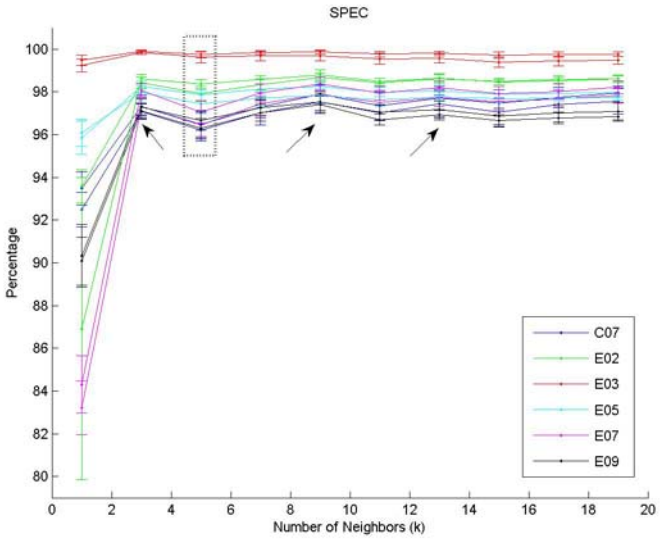
This experiment examines the performance of the k-NN rule in classifying the data of this research under varying values of the parameter k. The objective is to determine the most appropriate value of k for the classifier according to the metrics sensitivity and specificity.

Ten trials of 15 Monte Carlo simulations of tuning and testing a k-NN classifier are executed in which each simulation randomly generates a balanced training sample (50% of the values from the class for the epileptic oscillations and the remaining half of the sample for the complementary class) and an unbalanced testing sample in which 80% (20%) of the points represented the class for the epileptic oscillations (complementary class). In each trial, the number of neighbors is increased in odd numbered increments (e.g. k = 1, 3, 15). Odd numbers are used to avoid ties in the decision of the k-NN. In summary, k = 5 is deemed appropriate for the nearest neighbor rule for each subject because the sensitivity peaked at that value, while the specificity does not vary considerably for $k \geq 5$ in each subject. Figure F.2-4 illustrates the results of the abovementioned analysis.

The subsequent section further discusses the implications of a mismatch between the proportions of classes in the training and testing samples. The reason for the chosen proportions and disparity between random samples is to more closely simulate the problem of classifying rare events (epileptic oscillations) in a testing set. The 80%-20% division for the testing sample reflects the testing data that is constructed according to Section 3.1. A correction factor, which also is discussed in the next section, is assessed to compensate for the "mismatched" training and testing samples.



a



b

Figure F.2-4 This figure shows that five neighbors maximized the sensitivity (*a*) while not significantly sacrificing the specificity (*b*) of the *k*-NN classifier. The arrows in the bottom plot signify potential peaks in the curve for the specificity that correspond to considerably lower measures of sensitivity than the value at $k = 5$.

F.2.3 Mismatched Class Proportions in Training and Testing Samples

In general, training a classifier with one proportion of points in a sample and testing on data that actually exhibits a different proportion may lead to misleading or suboptimal metrics of performance if no correction to account for the mismatch is implemented [181]. The cause of this issue is that the apparent *a priori* probability distributions that are estimated from the proportion of distinct classes in the training sample usually will not reflect the distributions that should have been estimated from the true occurrence of the classes as observed during classification of the testing data. However, the report [181] devised a solution to this problem: train a classifier using balanced classes and adjust the classifier during testing according to Equation F.2-1, where $P_{true}(C_i)$ is the true proportion of the i^{th} class, in the testing sample and $P_{tune}(C_i)$ is the known proportion of the i^{th} class in the training sample, C is the number of distinct classes, and the **v** is a row vector representing the scores of the classifier before entering its competitive layer. In the case of dichotomous classes, as with this research, Equation F.2-1 becomes Equation F.2-2. For experiments such as Section F.2.2, Equation F.2-3 was applied, where **v** contained the votes for each binary class among the *k* nearest neighbors before a majority vote decided the class of each point in the testing sample.

$$correction = \mathbf{v} \bullet [\frac{P_{true}(i)}{P_{tune}(i)}, \frac{P_{true}(i+1)}{P_{tune}(i+1)}, \cdots \frac{P_{true}(C)}{P_{tune}(C)}] \tag{F.2-1}$$

$$correction = \mathbf{v} \bullet [\frac{P_{true}(1)}{P_{tune}(1)} \; \frac{P_{true}(2)}{P_{tune}(2)}] \tag{F.2-2}$$

$$correction = \mathbf{v} \bullet [\frac{.8}{.5} \; \frac{.2}{.5}] = \mathbf{v} \bullet [1.6 \quad 0.4] \tag{F.2-3}$$

F.3 Frequent Itemset Mining using Other Physics-based Features

This section briefly illustrates the results of using a few c-features other than the *mean energy* for the spatial-temporal matrix of measures in the frequent itemset mining problem. Specifically, the feature mean energy is chosen because Worrell et al. demonstrates that the average high-frequency energy from epileptic brain (in the SOZ) was significantly greater ($p < 0.05$) than in the normal control region (brain outside of the SOZ) [53]; but the presented methodology accommodates alternative features. The results in the following figures illustrate that mean energy most consistently reveals clusters of channels coincident with epileptic networks within and across patients. The *mean curve length*, which is a feature that is correlated to mean energy, does exhibit some results similar to that of mining based upon a matrix of mean energy measures; but the remaining features (e.g., mean entropy, mean phase) demonstrates conflicting results.

As a final note, the problem of choosing a feature for the spatial-temporal matrix resembles the problem of choosing (creating) features for the detector of high frequency epileptiform oscillations. The latter task is readily implemented with an evolutionary algorithm because an appropriate objective function (separation of classes) can be defined; however, the former task is not resolved in the same manner because an apt objective function (e.g., separation of regions of epileptic networks from regions without epileptic networks) could not be defined until evidence is presented that reliably demonstrates the occurrence of concentrated HFEO's as indispensable for epileptic networks along with corresponding regions of those networks—which is the overall problem of this research.

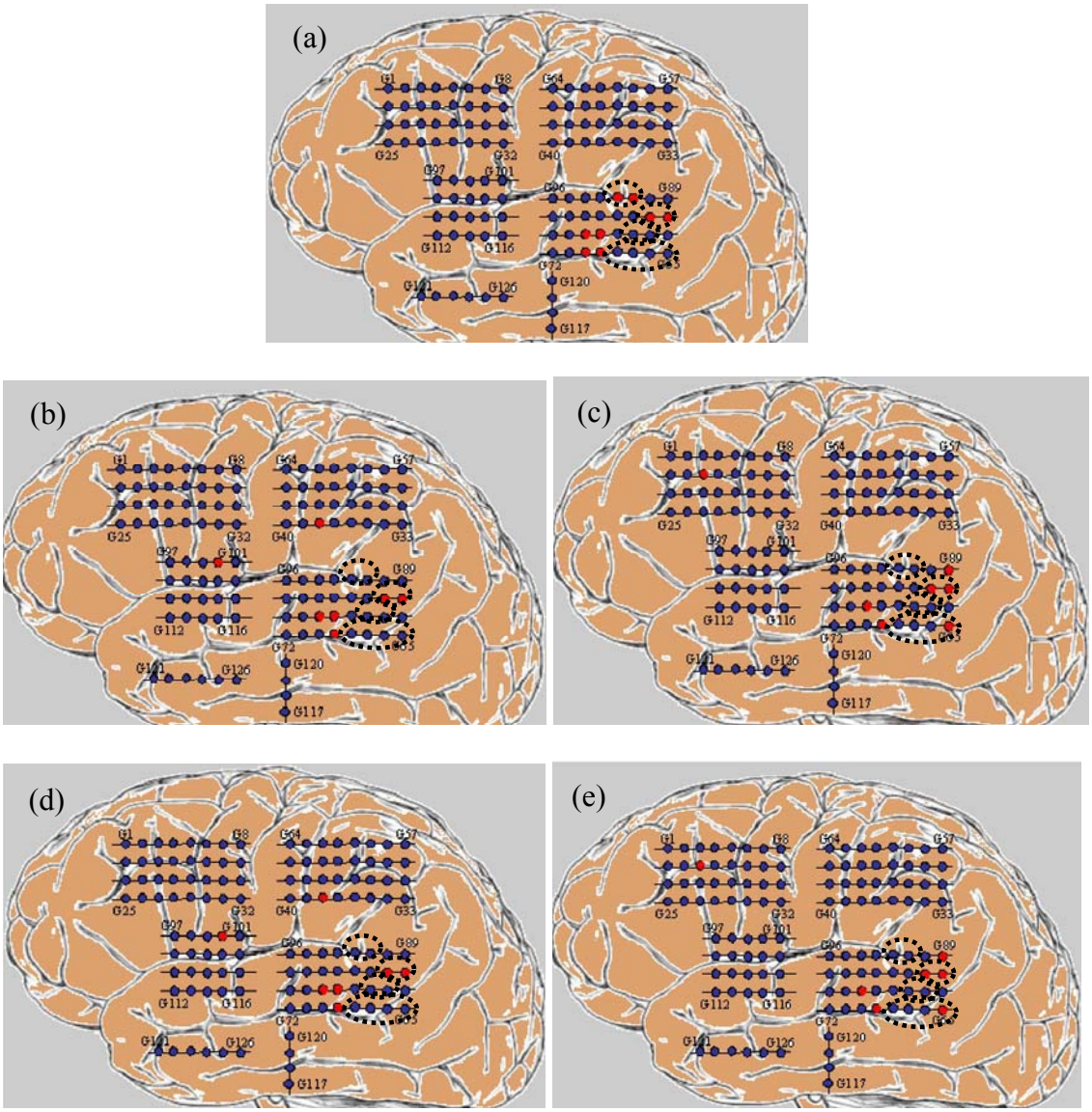


Figure F.3-1 This figure shows the results of FIM with a spatial-temporal matrix of measures using the following features as measures for subject C07: mean energy (*a*), mean curve length (*b*), mean entropy (*c*), mean frequency (*d*), and mean phase (*e*). The choice in feature does not appear to make a difference in the output of the FIM because the input reflects considerably great separation between regions with (*dashed ovals*) and without epileptic networks (c.f. Figure 4.3-6).

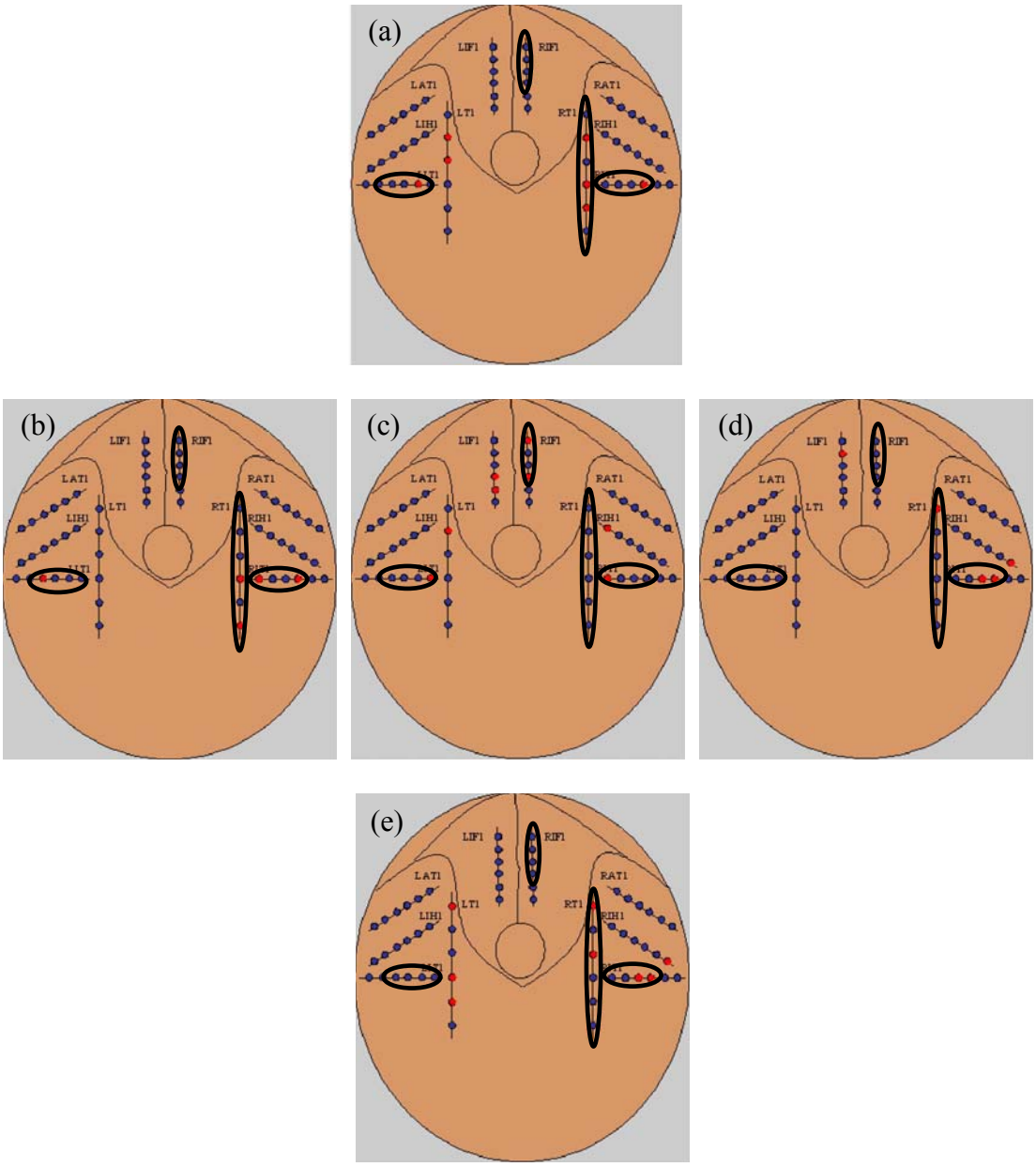


Figure F.3-2 This figure shows the results of FIM with a spatial-temporal matrix of measures using the following features as measures for subject E03: mean energy (*a*), mean curve length (*b*), mean entropy (*c*), mean frequency (*d*), and mean phase (*e*). The features mean energy and mean curve length coincide more with regions related to epileptic networks (*dashed ovals*).

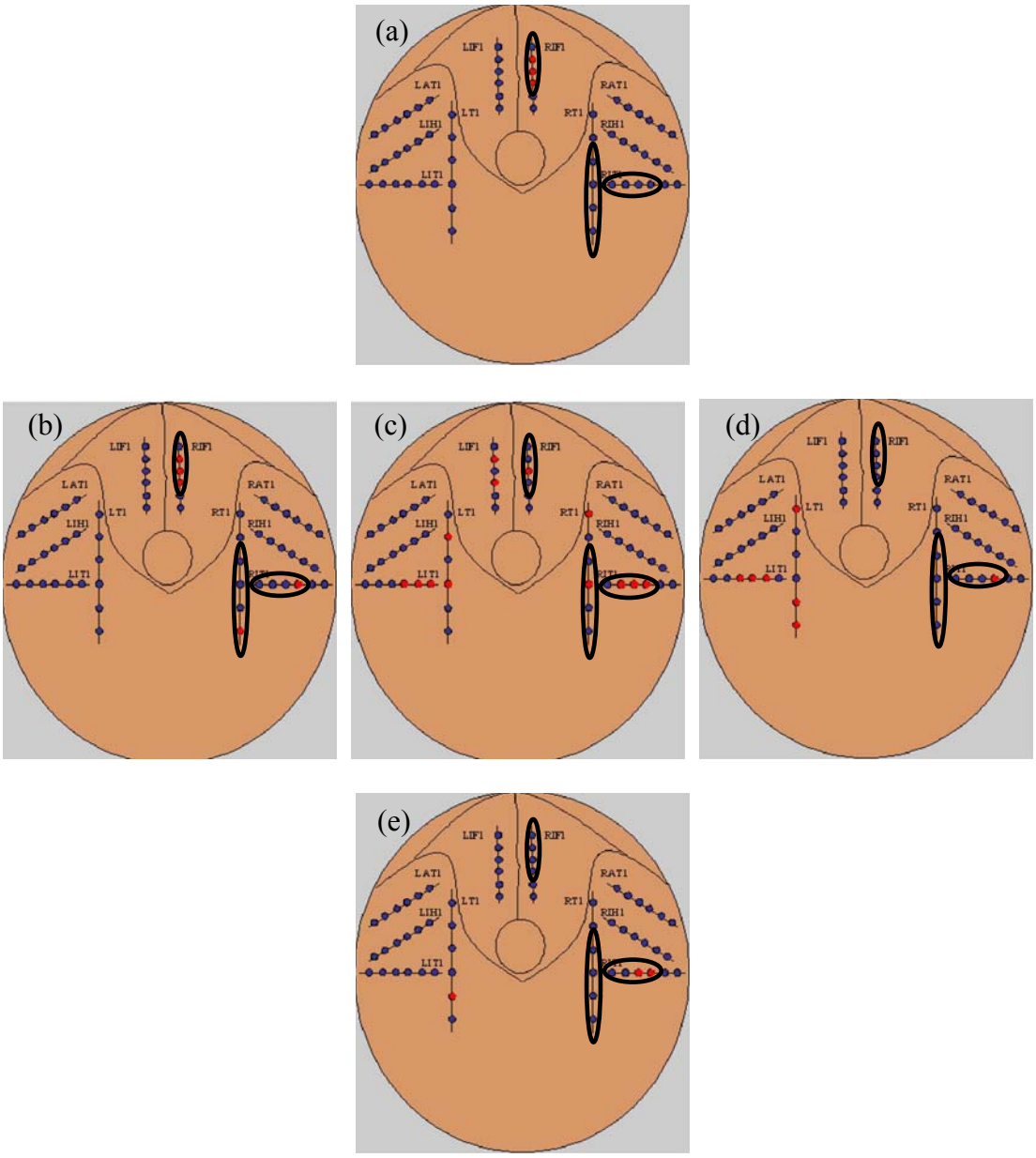


Figure F.3-3 This figure shows the results of FIM with a spatial-temporal matrix of measures using the following features as measures for subject E05: mean energy (*a*), mean curve length (*b*), mean entropy (*c*), mean frequency (*d*), and mean phase (*e*). The features mean energy and mean curve length coincide more with regions related to epileptic networks (*dashed ovals*).

References

[1] E. R. Kandel, J. H. Schwartz, and T. M. Jessel, *Principles of Neural Science*, 3rd ed: Prentice-Hall, 1991.

[2] E. F. o. America, *"Epilepsy Facts & Figures"*, www.efa.org, 2003.

[3] W. A. Hauser and D. C. Hesdorffer, *Epilepsy: frequency, causes, and consequences*. New York, 1990.

[4] T. N. Centre, *"The Epicentre"*, http:/137.172.248.46/epilepsy.htm, 1997.

[5] G. L. Holmes, "The Comprehensive Evaluation and Treatment of Epilepsy: A practical guide," in *Academic Press*, 1997.

[6] F. Drislane and D. Schomer, "Clinical implications of generalized electrographic status epilepticus," *Epilepsy Research*, vol. 19, no. 2, pp. 111-121, 1994.

[7] M. R. Mehta, C. Dasgupta, and G. R. Ulal, "Kindling of focal epilepsy occurs due to learning: a neural network model," *International Journal of Neural Systems*, vol. 6, pp. 107-111, 1995.

[8] P. L. Nunez, "Neocortical Dynamics and Human EEG Rhythms," in *Oxford University Press Inc.* New York, 1995.

[9] M. R. Isley et al., "Electromyography/Electroencephalography,", 1 ed. Washington, D. C.: Spacelabs Medical, 1993.

[10] R. D. Traub, G. R. Jefferys, and M. A. Whittington, "Fast Oscillations in Cortical Circuits," in *The MIT Press*. Cambridge, 1999.

[11] S. Wiebe, W. T. Blume, J. P. Girvin, and M. Eliasziw, "A randomized, controlled trial of surgery for temporal-lobe epilepsy," *N Engl J Med*, vol. 345, no. 5, pp. 311-8., 2001.

[12] C. S. Schachter and D. Schmidt, "Vagus Nerve Stimulation," 2001.

[13] J. Kerrigan et al., "Electrical Stimulation Of The Anterior Nucleus Of The Thalamus For The Treatment Of Intractable Epilepsy," *Epilepsia*, vol. *In press*, 2003.

[14] K. Vonck et al., "Long-term amygdalohippocampal stimulation for refractory temporal lobe epilepsy," *Ann Neurol*, vol. 52, no. 5, pp. 556-65., 2002.

[15] A. Umemura et al., "Deep brain stimulation for movement disorders: morbidity and mortality in 109 patients," *J Neurosurg*, vol. 98, no. 4, pp. 779-84., 2003.

[16] B. Pless, "Automatic, responsive brain stimulation in human epilepsy (report of 34 patients)," in *Proc. Spring Epilepsy Research Conference*. Georgetown, Grand Cayman, BWI, 2003. Conference presentation, not published.

[17] M. Hodaie, R. A. Wennberg, J. O. Dostrovsky, and A. M. Lozano, "Chronic anterior thalamus stimulation for intractable epilepsy," *Epilepsia*, vol. 43, no. 6, pp. 603-8, 2002.

[18] D. S. Dinner et al., "EEG and evoked potential recording from the subthalamic nucleus for deep brain stimulation of intractable epilepsy," *Clin Neurophysiol*, vol. 113, no. 9, pp. 1391-402., 2002.

[19] S. Chabardes et al., "Deep brain stimulation in epilepsy with particular reference to the subthalamic nucleus," *Epileptic Disorders*, vol. 4 Suppl 3, pp. S83-93, 2002.

[20] A. L. Benabid et al., "Antiepileptic effect of high-frequency stimulation of the subthalamic nucleus (corpus luysi) in a case of medically intractable epilepsy caused by focal dysplasia: a 30-month follow-up: technical case report," *Neurosurgery*, vol. 50, no. 6, pp. 1385-91; discussion 1391-2., 2002.

[21] F. Velasco et al., "Electrical Stimulation for Epilepsy: Stimulation of Hippocampal Foci," *Stereotact Funct Neurosurg*, vol. 77, no. 1-4, pp. 223-227., 2001.

[22] M. Velasco, F. Velasco, and A. L. Velasco, "Centromedian-thalamic and hippocampal electrical stimulation for the control of intractable epileptic seizures," *J Clin Neurophysiol*, vol. 18, no. 6, pp. 495-513., 2001.

[23] F. Velasco et al., "Stimulation of the central median thalamic nucleus for epilepsy," *Stereotact Funct Neurosurg*, vol. 77, no. 1-4, pp. 228-32, 2001.

[24] I. Osorio et al., "An introduction to contingent (closed-loop) brain electrical stimulation for seizure blockage, to ultra-short-term clinical trials, and to multidimensional statistical analysis of therapeutic efficacy," *Journal of Clinical Neurophysiology*, vol. 18, no. 6, pp. 533-44, 2001.

[25] T. Loddenkemper et al., "Deep brain stimulation in epilepsy," *J Clin Neurophysiol*, vol. 18, no. 6, pp. 514-32., 2001.

[26] B. Litt and G. Baltuch, "Brain Stimulation for Epilepsy," *Epilepsy, Brain and Behavior*, 2001.

[27] D. M. Durand and M. Bikson, "Suppression and control of epileptiform activity by electrical stimulation: a review," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1065-1082, 2001.

[28] M. Velasco et al., "Subacute electrical stimulation of the hippocampus blocks intractable temporal lobe seizures and paroxysmal EEG activities," *Epilepsia*, vol. 41, no. 2, pp. 158-69., 2000.

[29] R. S. Fisher and J. Ho, "Potential new methods for antiepileptic drug delivery," *CNS Drugs*, vol. 16, no. 9, pp. 579-93, 2002.

[30] A. G. Stein et al., "An automated drug delivery system for focal epilepsy," *Epilepsy Res*, vol. 39, no. 2, pp. 103-14, 2000.

[31] S. Ward and M. Rise, "Techniques for treating epilepsy by brain stimulation and drug infusion," US patent 5713923, Feb. 3, 1998.

[32] M. Musicco and F. S. T. Group, "Randomized clincal trial on the efficacy of antiepileptic drugs in reducing the risk of relapse after a first unprovoked tonic-clonic seizure," *Neurology*, vol. 43, no. March, pp. 478-483, 1993.

[33] C. Dodrill, "Problems in the Assessment of Cognitive Effects of Antiepileptic Drugs," *Epilepsia*, vol. 33(Suppl. 6), pp. S29-32, 1992.

[34] E. Vining, "Cognitive dysfunction Associated with Antiepileptic Drug Therapy," *Epilepsia*, vol. 28(Suppl. 2), pp. S18-22, 1987.

[35] M. Trimble, "Anticonvulsant drugs and cognitive function: a review of the literature," *Epilepsia*, vol. 28, pp. 37-45, 1987.

[36] C. B. Dodrill, "Effects of Antiepileptic Drugs on Abilities," *Journal of Clinical Psychiatry*, vol. 49, no. 4, Suppl, pp. 31-34, 1988.

[37] M. Trimble, "Antiepileptic Drugs, Cognitive function, and Behavior in Children: Evidence from Recent Studies," *Epilepsia*, vol. 31(Suppl 4), pp. S30-34, 1990.

[38] C. Valmadrid, C. Voorhees, B. Litt, and C. R. Schneyer, "Practice patterns of neurologists regarding bone and mineral effects of antiepileptic drug therapy. [see comments.]," *Archives of Neurology*, vol. 58, no. 9, pp. 1369-74, 2001.

[39] X. F. Yang and S. M. Rothman, "Focal cooling rapidly terminates experimental neocortical seizures," *Ann Neurol*, vol. 49, no. 6, pp. 721-6., 2001.

[40] X. F. Yang, D. W. Duffy, R. E. Morley, and S. M. Rothman, "Neocortical seizure termination by focal cooling: temperature dependence and automated seizure detection," *Epilepsia*, vol. 43, no. 3, pp. 240-5, 2002.

[41] K. M. Karkar et al., "Focal cooling suppresses spontaneous epileptiform activity without changing the cortical motor threshold," *Epilepsia*, vol. 43, no. 8, pp. 932-5, 2002.

[42] M. W. Hill, M. Wong, A. Amarakone, and S. M. Rothman, "Rapid cooling aborts seizure-like activity in rodent hippocampal- entorhinal slices," *Epilepsia*, vol. 41, no. 10, pp. 1241-8., 2000.

[43] C. A. Tassinari, M. Cincotta, G. Zaccara, and R. Michelucci, "Transcranial magnetic stimulation and epilepsy," *Clin Neurophysiol*, vol. 114, no. 5, pp. 777-798., 2003.

[44] R. S. Ghai, M. Bikson, and D. M. Durand, "Effects of applied electric fields on low-calcium epileptiform activity in the CA1 region of rat hippocampal slices," *J Neurophysiol*, vol. 84, no. 1, pp. 274-80, 2000.

[45] K. A. Richardson et al., "In vivo modulation of hippocampal epileptiform activity with radial electric fields," *Epilepsia*, vol. 44, no. 6, pp. 768-77, 2003.

[46] B. Gluckman, H. Nguyen, S. Weinstein, and S. Schiff, "Adaptive Electric Field Control of Epileptic Seizures," *Journal of Neuroscience*, vol. 21, no. 2, pp. 590-600, 2001.

[47] M. Bikson et al., "Suppression of epileptiform activity by high frequency sinusoidal fields in rat hippocampal slices," *J Physiol*, vol. 531, no. Pt 1, pp. 181-91., 2001.

[48] H. H. Jasper, "Physiopathological Mechanisms of Post-traumatic Epilepsy," *Epilepsia*, vol. 11, pp. 73-80, 1970.

[49] J. Engel, ed. *Surgical Treatment of the Epilepsies*, 1st ed. New York: Raven Press, 1987, vol. 1.

[50] L. F. Quesney, "Intracranial EEG Investigation in Neocorical Epilepsy," *Adv Neurol*, vol. 84, pp. 253-74, 2000.

[51] L. F. Quesney et al., "Presurgical EEG Investigation in Frontal Lobe Epilepsy," *Epilepsy Research Supplement*, vol. 5, pp. 55-69, 1992.

[52] B. Litt and G. W. Worrell, personal communication, "The Use and Placement of IEEG Electrodes for Human Subjects," 2004.

[53] G. A. Worrell et al., "High Frequency Oscillations and Seizure Generation in Neocortical Epilepsy," *Brain*, vol. 127, pp. 1-11, April 2004.

[54] H. Hassanpour and M. Mesbah, "Neonatal EEG seizure detection using spike signatures in the time-frequency domain," in *Proc. Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, 2003, vol. 2, pp. 41-44 vol.2.

[55] R. D. A. Esteller, M; Echauz, J; Vachtsevanos, G; Litt, B, "Intelligent Approach for Detecting Electrographic Seizure Onset in Patients with Mesial Temporal Lobe Epilepsy," *Annals of Biomedical Engineering*, vol. in revision, 2003.

[56] Esteller, R., Echauz, J., M. D'Alessandro, G. Vachtsevanos, B. Litt, "Real-Time Simulation of a Seizure Detection System Suitable for an Implantable Device," in *Proc. American Epilepsy Society*. Philadelphia, 2002.

[57] T. E. Peters, N. C. Bhavaraju, M. G. Frei, and I. Osorio, "Network system for automated seizure detection and contingent delivery of therapy," *Journal of Clinical Neurophysiology*, vol. 18, no. 6, pp. 545-9, 2001.

[58] K. K. Jerger et al., "Early seizure detection," *J Clin Neurophysiol*, vol. 18, no. 3, pp. 259-68., 2001.

[59] R. E. Esteller, J.; D'Alessandro, M.; Vachtsevanos, G.; Litt, B., "Feature parameter optimization for seizure detection/prediction," in *Proc. 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2001, vol. 2, pp. 1711-1714.

[60] R. E. Esteller, J.; Tcheng, T.; Litt, B.; Pless, B., "Line length: an efficient feature for seizure onset detection," in *Proc. 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2001, vol. 2, pp. 1707 -1710.

[61] C. J. James and D. Lowe, "Using independent component analysis & dynamical embedding to isolate seizure activity in the EEG," in *Proc.*

Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE, 2000, vol. 2, pp. 1329-1332 vol.2.

[62] R. Esteller, *Detection of Seizure Onset in Epileptic Patients form Intracranial EEG Signals*, Georgia Institute of Technology, 2000.

[63] R. Esteller et al., "Fractal Dimension Detects Seizure Onset in Mesial Temporal Lobe Epilepsy," in *Proc. IEEE International Conference on Engineering in Medicine and Biology*. Atlanta, Georgia, 1999.

[64] J. Echauz et al., "Median-Based Filtering Methods for EEG Seizure Detection," in *Proc. IEEE International Conference on Engineering in Medicine and Biology*. Atlanta, Georgia, 1999.

[65] I. Osorio, M. Frei, and S. Wilkinson, "Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset," *Epilepsia*, vol. 39, no. 6, pp. 615-627, 1998.

[66] L. Hively, N. Clapp, C. S. Daw, and W. Lawkins, "Apparatus and method for epileptic seizure detection using non-linear techniques," US patent 5743860, Apr. 28, 1998.

[67] H. Qu and J. Gotman, "A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: Possible use as a warning device," *IEEE Transactions of Biomedical Engineering*, vol. 44, no. 2, pp. 115-122, Feb. 1997.

[68] A. Gabor, R. Leach, and R. Dowla, "Automated seizure detection using a self-organizing neural network," *Electroencephalography and Clinical Neurophysiology*, vol. 9, pp. 257-266, 1996.

[69] I. Osorio et al., "A Method for Accurate Automated Real-Time Seizure Detection," *Epilepsia*, vol. 36, no. suppl. 4, 1995.

[70] I. Osorio, M. Frei, D. Lerner, and S. Wilkinson, "Automated rapid seizure detection in the human ECoG," in *Proc. Eighth IEEE Symposium on Computer-Based Medical Systems*, June 9-10, 1995, pp. 328.

[71] J. Gotman, J. Zhang, B. Rosenblatt, and R. Gottesman, "Automatic seizure detection in newborns and infants," in *Proc. Engineering in Medicine and Biology Society, 1995. IEEE 17th Annual Conference*, 1995, vol. 2, pp. 913-914 vol.2.

[72] J. Echauz and G. Vachtsevanos, "Neural network detection of antiepileptic drugs from a single EEG trace," in *Proc. Electro 94 International Conference*. Boston, MA, May 10-12, 1994, pp. 346-351.

[73] M. Bozek-Kuzmicki, D. Colella, and G. M. Jacyna, "Feature-based epileptic seizure detection and prediction from ECoG recordings," in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*. Philadelphia, PA, Oct. 25-28, 1994, pp. 564-567.

[74] A. B. Gardner, *A Novelty Detection Approach to Seizure Analysis from Intracranial EEG*, Georgia Institute of Technology, 2004.

[75] M. D'Alessandro, *The Utility of intracranial EEG Feature and Channel Synergy for Evaluating the Spatial and Temporal Behavior of Oscillations*. Ph.D. Dissertation, Georgia Institute of Technology, 2001.

[76] Echauz, J., Esteller, R., Litt, B. and Vachtsevanos, G., "Unified Probabilistic Framework for Predicting and Detecting Seizure Onsets in the Brian and Multitherapeutic Device", U. S. Patent No. 6,678,548, issued January 13, 2004.

[77] J. Echauz, B. Litt, R. Esteller, and G. Vachtsevanos, "Three Decades of Engineered Methods and Systems for Predicting Epileptic Seizures," *Georgia Institute of Technology Presentation*, September, 1999.

[78] L. D. Iasemidis, "Epileptic Seizure Prediction and Control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 549-558, May 2003.

[79] L. D. Iasemidis et al., "Adaptive epileptic seizure prediction system," *IEEE Trans Biomed Eng*, vol. 50, no. 5, pp. 616-27., 2003.

[80] Y. C. Lai, M. A. Harrison, M. G. Frei, and I. Osorio, "Controlled test for predictive power of Lyapunov exponents: their inability to predict epileptic seizures," *Chaos*, vol. 14, no. 3, pp. 630-42, Sep 2004.

[81] M. Le Van Quyen et al., "Anticipation of epileptic seizures from standard EEG recordings," *The Lancet*, vol. 357, no. 9251, pp. 183, 2001.

[82] K. Lehnertz et al., "Forecasting epileptic seizures: Adequacy of different EEG analysis techniques," *Epilepsia*, vol. 41, no. suppl. 7, pp. 212 (abstract), Dec. 2000.

[83] K. Lehnertz and C. Elger, "Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity," *Physical Review Letters*, vol. 80, no. 22, pp. 5019-5022, 1998.

[84] K. Lehnertz et al., "Seizure prediction by nonlinear EEG analysis," *IEEE Eng Med Biol Mag*, vol. 22, no. 1, pp. 57-63., 2003.

[85] K. Lehnertz et al., "Is it possible to anticipate seizure onset by non-linear analysis of intracerebral EEG in human partial epilepsies?," *Rev Neurol (Paris)*, vol. 155, no. 6-7, pp. 454-6, 1999.

[86] B. Litt and J. Echauz, "Prediction of Epileptic Seizures," *The Lancet Neurology*, vol. 1, no. 1, pp. 22-30, May 2002 2002.

[87] B. Litt et al., "Translating seizure detection, prediction and brain stimulation into implantable devices for epilepsy," in *Proc. Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, 2003, pp. 485-488.

[88] M. McKeown and J. O. McNamara, "When Do Epileptic Seizures *Really* Begin?," *Neuron*, vol. 30, no. 1-9, pp. 1-3, April 2001.

[89] P. E. McSharry, L. A. Smith, and L. Tarassenko, "Comparison of predictability of epileptic seizures by a linear and a nonlinear method," *IEEE Trans Biomed Eng*, vol. 50, no. 5, pp. 628-33., 2003.

[90] J. S. Paul et al., "Prediction of PTZ-induced seizures using wavelet-based residual entropy of cortical and subcortical field potentials," *IEEE Trans Biomed Eng*, vol. 50, no. 5, pp. 640-8., 2003.

[91] I. Osorio, M. A. Harrison, Y. C. Lai, and M. G. Frei, "Observations on the application of the correlation dimension and correlation integral to the prediction of seizures," *Journal of Clinical Neurophysiology*, vol. 18, no. 3, pp. 269-74, 2001.

[92] A. Petrosian, R. Homan, S. Pemmaraju, and S. Mitra, "Wavelet-based Texture Analysis of EEG Signal for Prediction of Epileptic Seizure," in *Proc. SPIE Proceedings on Wavelet Applications in Signal and Image Processing*. San Diego, CA, July 12-14, 1995, vol. 2569, pp. 189-194.

[93] S. Sabesan et al., "Predictability of epileptic seizures: a comparative study using Lyapunov exponent and entropy based measures," *Biomedical Sciences Instrumentation*, vol. 39, pp. 129-35, 2003.

[94] R. Venugopal et al., "A new approach towards predictability of epileptic seizures: KLT dimension," *Biomedical Sciences Instrumentation*, vol. 39, pp. 123-8, 2003.

[95] H. Witte, L. D. Iasemidis, and B. Litt, "Special Issue on Epileptic Seizure Prediction," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 537-539, May 2003.

[96] M. Winterhalder et al., "The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods," *Epilepsy Behav*, vol. 4, no. 3, pp. 318-25, Jun 2003.

[97] P. Rajna and e. al, "Hungarian multicentre epidemiologic study of the warning and initial symptoms (prodrome, aura) of epileptic seizures," *Seizure*, vol. 6, pp. 361-368, Oct. 1997.

[98] B. Litt et al., "Epileptic Seizures May Begin Hours in Advance of Clinical Onset: A Report of Five Patients," *Neuron*, vol. Vol. 30, pp. 51–64, April 2001 2001.

[99] C. Baumgartner et al., "Preictal SPECT in temporal lobe epilepsy: Regional cerebral blood flow is increased prior to electroencephalography-seizure onset," *Journal of Nuclear Medicine*, vol. 39, no. 6, pp. 978-982, June 1998.

[100] M. Weinand et al., "Cerebral blood flow and temporal lobe epileptogenicity," *Journal of Neurosurgery*, vol. 86, no. 2, pp. 226-232, Feb. 1997.

[101] B. Westmoreland, D. Klass, and F. Sharbrough, "Chronic periodic lateralized epileptiform discharges," *Archives of Neurology*, vol. 43, pp. 494-496, 1986.

[102] D. Contreras, A. Destexhe, T. J. Sejnowski, and M. Steriade, "Spatiotemporal patterns of spindle oscillations in cortex and thalamus," *Journal of Neuroscience*, vol. 17, no. 3, pp. 1179-96, 1997.

[103] D. Contreras, R. Curro Dossi, and M. Steriade, "Bursting and tonic discharges in two classes of reticular thalamic neurons," *Journal of Neurophysiology*, vol. 68, no. 3, pp. 973-7, 1992.

[104] M. Chavez et al., "Spatio-temporal dynamics prior to neocortical seizures: amplitude versus phase couplings," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 5, pp. 571-583, 2003.

[105] A. W. Bragin, C., Almajano, J., Mody, I., Engel Jr., J., "High Frequency Oscillations after Status Epilepticus: Epileptogenesis and Seizure Genesis," *Epilepsia*, vol. 45, no. 9, pp. 1017-1023, 2004.

[106] A. Bragin et al., "Interictal high-frequency oscillations (80-500 Hz) in the human epileptic brain: entorhinal cortex," *Annals of Neurology*, vol. 52, no. 4, pp. 407-15, 2002.

[107] A. Bragin, I. Mody, C. L. Wilson, and J. Engel, Jr., "Local generation of fast ripples in epileptic brain," *Journal of Neuroscience*, vol. 22, no. 5, pp. 2012-21, 2002.

[108] A. Bragin et al., "Hippocampal and entorhinal cortex high-frequency oscillations (100--500 Hz) in human epileptic brain and in kainic acid--treated rats with chronic seizures," *Epilepsia*, vol. 40, no. 2, pp. 127-37, 1999.

[109] A. Bragin et al., "High-frequency oscillations in human brain," *Hippocampus*, vol. 9, no. 2, pp. 137-42, 1999.

[110] W. T. Blume, A. G. Parrent, and M. Kaibara, "Stereotactic amygdalohippocampotomy and mesial temporal spikes," *Epilepsia*, vol. 38, no. 8, pp. 930-6, 1997.

[111] S. D. Cranstoun, G. Worrell, J. Echauz, and B. Litt, "Self-organized criticality in the epileptic brain," in *Proc. [Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, 2002, vol. 1, pp. 232-233 vol.1.

[112] M. Ge et al., "Numeral research of synchronous oscillation on two Chay neurons and 2-D network by gap junction," in *Proc. Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, 2003, vol. 2, pp. 1913-1916 Vol.2.

[113] J. Gotman, "Relationships between interictal spiking and seizures: Human and experimental evidence," *Canadian Journal of Neurologic Sciences*, vol. 18, no. Suppl. 4, pp. 573-576, Nov. 1991.

[114] E. Halgren and C. Wilson, "Recall deficits produced by after discharges in the human hippocampal formation and amygdala," *Electroencephalography and Clinical Neurophysiology*, vol. 61, pp. 375-80, 1985.

[115] C. Huang and L. J. White, "High-frequency components in epileptiform EEG," *Journal of Neuroscience Methods*, vol. 30, no. 3, pp. 197-201, 1989.

[116] P. Kahane et al., "An H(2) (15)O-PET study of cerebral blood flow changes during focal epileptic discharges induced by intracerebral electrical stimulation," *Brain*, vol. 122, no. Pt 10, pp. 1851-65, 1999.

[117] M. Isokawa-Akesson, C. Wilson, and T. Babb, "Diversity in periodic pattern of firing in human hippocampal neurons," *Experimental Neurology*, vol. 98, pp. 137-51, 1987.

[118] A. Katz, D. Marks, G. McCarthy, and S. Spencer, "Does interictal spiking change prior to seizures?," *Electroencephalography and Clinical Neurophysiology*, vol. 79, no. 2, pp. 153-156, Aug. 1991.

[119] M. Le Van Quyen et al., "Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures," *Eur J Neurosci*, vol. 12, no. 6, pp. 2124-34, 2000.

[120] B. Litt et al., "Pre-Ictal Prodromes Predict Seizures in a Patient With Mesial Temporal Lobe Epilepsy.," in *Proc. American Epilepsy Society Meeting*. Orlando, Florida, December, 1999.

[121] F. Matsuo, "Trajectory analysis in human epileptiform discharges of temporal lobe origin," in *Proc. Neural Networks, 1990., 1990 IJCNN International Joint Conference on*, 1990, pp. 753-758 vol.2.

[122] G. Mayer-Kress et al., "Dimensional analysis of nonlinear oscillations in brain, heart, and muscle," *Mathematical Biosciences*, vol. 90, no. 1/2, pp. 155-182, 1988.

[123] F. Mormann, K. Lehnertz, R. Andrzejak, and C. Elger, "Characterizing preictal states by changes in phase synchronization in intracranial EEG recordings from epilepsy patients," *Epilepsia*, vol. 41, no. suppl. 7, pp. 167 (abstract), Dec. 2000.

[124] J. Reiher, J. Rivest, F. Grand'Maison, and C. Leduc, "Periodic lateralized epileptiform discharges with transitional rhythmic discharges: association with seizures," *Electroencephalography and Clinical Neurophysiology*, vol. 78, pp. 12-17, 1991.

[125] M. Sperling and M. O'Connor, "Auras and Subclinical Seizures: Characteristics and Prognostic Significance," *Annals of Neurology*, vol. 28, pp. 320-328, 1990.

[126] M. Steriade and D. Contreras, "Spike-wave complexes and fast components of cortically generated seizures. I. Role of neocortex and thalamus," *Journal of Neurophysiology*, vol. 80, no. 3, pp. 1439-55, 1998.

[127] G. A. Worrell et al., "Focal ictal beta discharge on scalp EEG predicts excellent outcome of frontal lobe epilepsy surgery," *Epilepsia*, vol. 43, no. 3, pp. 277-82., 2002.

[128] M. Steriade, D. Contreras, F. Amzica, and I. Timofeev, "Synchronization of fast (30-40 Hz) spontaneous oscillations in intrathalamic and thalamocortical networks," *Journal of Neuroscience*, vol. 16, no. 8, pp. 2788-808, 1996.

[129] M. Steriade, D. Contreras, and F. Amzica, "Synchronized sleep oscillations and their paroxysmal developments," *Trends in Neurosciences*, vol. 17, no. 5, pp. 199-208, 1994.

[130] K. Wilson et al., "Detection of epileptiform spikes in the EEG using a patient-independent neural network," in *Proc. Computer-Based Medical Systems, 1991. Proceedings of the Fourth Annual IEEE Symposium*, 1991, pp. 264-271.

[131] G. Van Hoey et al., "Combined detection and source analysis of epileptic EEG spikes," in *Proc. Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, 1998, vol. 4, pp. 2159-2162 vol.4.

[132] C.-W. Ko, Y.-D. Lin, H.-W. Chung, and G.-J. Jan, "An EEG spike detection algorithm using artificial neural network with multi-channel correlation," in *Proc. Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, 1998, vol. 4, pp. 2070-2073 vol.4.

[133] L. Tarassenko, Y. U. Khan, and M. R. G. Holt, "Identification of inter-ictal spikes in the EEG using neural network analysis," *Science, Measurement and Technology, IEE Proceedings*-, vol. 145, no. 6, pp. 270-278, 1998.

[134] G. Calvagno, M. Ermani, R. Rinaldo, and F. Sartoretto, "A multiresolution approach to spike detection in EEG," in *Proc. Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, 2000, vol. 6, pp. 3582-3585 vol.6.

[135] L.-s. Pon, M. Sun, and R. J. Sclabassi, "The bi-directional spike detection in EEG using mathematical morphology and wavelet transform," in *Proc. Signal Processing, 2002 6th International Conference on*, 2002, vol. 2, pp. 1512-1515 vol.2.

[136] R. J. Staba et al., "Quantitative analysis of high-frequency oscillations (80-500 Hz) recorded in human epileptic hippocampus and entorhinal cortex," *Journal of Neurophysiology*, vol. 88, no. 4, pp. 1743-52, 2002.

[137] M. Adjouadi et al., "Interictal spike detection using the Walsh transform," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 5, pp. 868-872, 2004.

[138] H. Hassanpour, M. Mesbah, and B. Boashash, "EEG spike detection using time-frequency signal analysis," in *Proc. Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, 2004, vol. 5, pp. V-421-V-424 vol.5.

[139] J. J. Niederhauser et al., "Detection of oscillations from depth-EEG using a sign periodogram transform," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 4, pp. 449-458, 2003.

[140] P. Fischer and R. Tetzlaff, "Pattern detection by cellular neuronal networks (CNN) in long-term recordings of a brain electrical activity in epilepsy," in *Proc. Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, 2004, vol. 1, pp. 169.

[141] P. Van Hese et al., "A new method for detection and source analysis of EEG spikes," in *Proc. Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, 2003, vol. 3, pp. 2455-2458 Vol.3.

[142] O. L. Smart, G. A. Worrell, B. Litt, and G. J. Vachtsevanos, "Automatic Detection of High Frequency Epileptiform Oscillations from the Intracranial EEG of Patients with Neocortical Epilepsy," in *Proc. 2005 Technical, Professional and Student Development Workshop (TPS)*. Boulder, CO, April 2005, 2005.

[143] O. L. Smart, Worrell, G. A., Litt, B., Vachtsevanos, G. J., "Automatic Detection Of EHFO From Focal Intracranial EEG Using Fuzzy C-Means Clustering," in *Proc. Biomedical Engineering: New Challenges for the Future. Proceedings of the 2004 Annual Fall Meeting (BMES)*. Philadelphia, PA, 15 October 2004, 2004.

[144] T. Zhang, F. Yang, and Q. Tang, "Wavelet based approach for detecting and classifying epileptiform waves in EEG," in *Proc. [Engineering in Medicine and Biology, 1999. 21st Annual Conf. and the 1999 Annual Fall Meeting of the Biomedical Engineering Soc.] BMES/EMBS Conference, 1999. Proceedings of the First Joint*, 1999, vol. 2, pp. 943 vol.2.

[145] R. D. Jones et al., "A system for detecting epileptiform discharges in the EEG: real-time operation and clinical trial," in *Proc. Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE*, 1996, vol. 3, pp. 948-949 vol.3.

[146] A. Ossadtchi et al., "Automated interictal spike detection and source localization in MEG using ICA and spatial-temporal clustering," in *Proc. Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*, 2002, pp. 785-788.

[147] J. R. Koza, "Hierarchical genetic algorithms operating on populations of computer programs," in *Proc. Proceedings of the 11th International Joint Conference on Artificial Intelligence*, 1989, vol. 1, pp. 768-774.

[148] J. R. Koza, "Genetic Programming: On Programming of Computers by Means of Natural Selection," in *MIT Press*. Cambridge, MA, 1992.

[149] J. R. Koza, "Genetic Programming II: Automatic Discovery of Reusable Programs," in *MIT Press*. Cambridge, MA, 1994.

[150] E. Marsh, personal communication, "The Observed Bandwidth of High-frequency Epiletiform Oscillations in Pediatric IEEG's,", 2005.

[151] R. Bellotti, M. Castellano, and F. De Carlo, "A chaotic map algorithm for knowledge discovery in time series: a case study on biomedical signals," *Nuclear Science, IEEE Transactions on*, vol. 51, no. 3, pp. 553-557, 2004.

[152] C. Bigan, "A system for neural networks detection and automatic identification of EEG epileptic events," in *Proc. Intelligent Methods in Healthcare and Medical Applications (Digest No. 1998/514), IEE Colloquium on*, 1998, pp. 13/1-13/4.

[153] F. Bodon, *APRIORI implementation of Ferenc Bodon*, http://www.cs.bme.hu/~bodon/en/apriori/.

[154] F. Bodon, *Apriori implementation for mining frequent sequence of items*, http://www.cs.bme.hu/~bodon/en/apriori_seq/index.html.

[155] F. Bodon, "A fast APRIORI implementation," in *Proc. CEUR Workshop Proceedings*. Melbourne, Florida, 2003. IEEE ICDM Workshop on Frequent Itemset Mining Implementations (FIMI'03), vol. 90.

[156] F. Bodon, "Surprising results of trie-based FIM algorithms," in *Proc. CEUR Workshop Proceedings*. Brighton, UK, November 2004, 2004. IEEE ICDM Workshop on Frequent Itemset Mining Implementations (FIMI'04), vol. 126.

[157] F. Bodon, "Trie-based Apriori Implementation for Mining Frequent Itemsequences," (B. G. a. S. N. a. M. J. Zaki, ed.). Chicago, IL, 2005. SIGKDD Workshop on Open Source Data Mining Workshop (OSDM'05).

[158] J. Bourien et al., "Mining reproducible activation patterns in epileptic intracerebral EEG signals: application to interictal activity," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 2, pp. 304-315, 2004.

[159] T. P. Exarchos et al., "A data mining based approach for the EEG transient event detection and classification," in *Proc. Computer-Based Medical Systems, 2005. Proceedings. 18th IEEE Symposium on*, 2005, pp. 35-40.

[160] Haykin, S. *Neural Networks: A comprehensive foundation*. 2nd ed. Upper Saddle River, NJ, Prentice Hall, 1999.

[161] Kennedy, J. and R. C. Eberhart. *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.

[162] Kennedy, J. and R. C. Eberhart, Particle Swarm Optimization. *Proceedings IEEE International Conference on Neural Networks*. 1995 4:1942-1948.

[163] Jirsch, J. D., E. Urrestarazu, P. LeVan, A. Oliver, F. Dubeau, and J. Gotman. High-frequency oscillations during human focal seizure. *Brain*. 2006. 129(6):1593-1608.

[164] Wikipedia contributors. Permutation test [Internet]. Wikipedia, The Free Encyclopedia; 2006 Jan 18, 19:06 UTC [cited 2006 Dec 12]. Available from: http://en.wikipedia.org/w/index.php?title=Permutation_test&oldid=35708137.

[165] Eberhart, R. C., Shi, Y. Comparison between Genetic Algorithms and Particle Swarm Optimization. Proc. Of the 7th Annual Conference on Evolutionary Programming. Berlin: Springer; 1998., pp. 611-618.

[166] R. Duda, P. Hart, and D. Stork, *Pattern Classification (2nd Edition)*: Wiley-Interscience, 2000.

[167] S. Fine, Y. Singer, and N. Tishby, "The Hierarchical Hidden Markov Model: Analysis and Applications," *Mach. Learn.*, vol. 32, no. 1, pp. 41-62, 1998/7// 1998.

[168] Wikipedia contributors. Hidden Markov model [Internet]. Wikipedia, The Free Encyclopedia; 2006 Dec 11, 02:26 UTC [cited 2006 Dec 12]. Available from: http://en.wikipedia.org/w/index.php?title=Hidden_Markov_model&oldid=93499942.

[169] W. D. Penny and S. J. Roberts, "Gaussian Observation Hidden Markov Models for EEG Analysis," Neural System Research Group, Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London, Technical Report TR-98-12, October 5, 1998 1998.

[170] G. L. Gerstein and B. Mandelbrot, "Random Walk Models for the Spike Activity of a Single Neuron," *Biophysical Journal*, vol. 4(1 Pt 1), pp. 41-68, January 1964 1964.

[171] S. Iyengar, "Hitting Lines With Two-Dimensional Brownian Motion," *Society for Industrial and Applied Mathematics*, vol. 45, no. 6, pp. 983-989, December 1985 1985.

[172] Wikipedia contributors. Brownian motion [Internet]. Wikipedia, The Free Encyclopedia; 2006 Dec 14, 23:45 UTC [cited 2006 Dec 15]. Available from: http://en.wikipedia.org/w/index.php?title=Brownian_motion&oldid=94396528.

[173] Wikipedia contributors. Inverse Gaussian distribution [Internet]. Wikipedia, The Free Encyclopedia; 2006 Nov 22, 06:09 UTC [cited 2006 Dec 15]. Available from:http://en.wikipedia.org/w/index.php?title=Inverse_Gaussian_distribution&oldid=89395952.

[174] F. Rosenow and H. Luders, "Presurgical evaluation of epilepsy 10.1093/brain/124.9.1683," *Brain*, vol. 124, no. 9, pp. 1683-1700, September 1, 2001 2001.

[175] Wikipedia contributors. Artificial neural network [Internet]. Wikipedia, The Free Encyclopedia; 2006 Dec 27, 05:44 UTC [cited 2006 Dec 30]. Available from: http://en.wikipedia.org/w/index.php?title=Artificial_neural_network&oldid=96690629.

[176] Shi, Y. and E. C. Eberhart, Parameter Selection in Particle Swarm Optimization. *Evolutionary Programming VII: Lectures Notes in Computer Science*. 1998 591-600.

[177] Wikipedia contributors. Central limit theorem [Internet]. Wikipedia, The Free Encyclopedia; 2007 Feb 1, 05:36 UTC [cited 2007 Feb 5]. Available from: http://en.wikipedia.org/w/index.php?title=Central_limit_theorem&oldid=104779331.

[178] Wikipedia contributors. Confidence interval [Internet]. Wikipedia, The Free Encyclopedia; 2007 Feb 3, 17:24 UTC [cited 2007 Feb 5]. Available from: http://en.wikipedia.org/w/index.php?title=Confidence_interval&oldid=105364113.

[179] J. L. Devore, Probability and Statistics for Engineering and the Sciences, 5th Edition: Duxbury/Thomson Learning, 2000.

[180] H. A. Firpi and E. D. Goodman, "Designing Templates for Cellular Neural Networks Using Particle Swarm Optimization," in *Proc. Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop (AIPR'04)*, 2004. IEEE Computer Society, vol. 00, pp. 119-123.

[181] J. Echauz, "Classifier-Based Performance Metrics and the Effects of Prior Probability Mismatches," IntelliMedix, Inc., Technical Report, March 2000.

[182] S. Haykin, *Neural Networks: A comprehensive foundation*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.

[183] Kil, D. H. and Shin, F. B. Pattern Recognition and Predictions with to Applications to Signal Characterization. AIP Press, 1996.

[184] M. L. Raymer et al., "Dimensionality Reduction Using Genetic Algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 4, no. 2, pp. 164-171, July 2000.